INTERNATIONAL SERIES IN OPERATIONS RESEARCH AND MANAGEMENT SCIENCE

Dynamic Optimization and Differential Games

Terry L. Friesz

Springer

VISIT…

International Series in Operations Research & Management Science

Volume 135

Series Editor

Frederick S. Hillier
Stanford University, CA, USA

Special Editorial Consultant

Camille C. Price
Stephen F. Austin State University, TX, USA

For other titles published in this series, go to
http://www.springer.com/series/6161

Terry L. Friesz

Dynamic Optimization and Differential Games

Springer

Terry L. Friesz
Pennsylvania State University
Dept. Industrial & Manufacturing Engineering
305 Leonhard Building
University Park, Pennsylvania
16802
USA
tfriesz@psu.edu

ISBN 978-0-387-72777-6 e-ISBN 978-0-387-72778-3
DOI 10.1007/978-0-387-72778-3
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: xxxxxxxxxx

© Springer Science+Business Media, LLC 2010
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Preface xiii

1 Introduction 1
1.1 Brief History of the Calculus of Variations and Optimal Control 2
1.2 The Brachistochrone Problem 4
1.3 Optimal Economic Growth 5
1.3.1 Ramsey's 1928 Model 5
1.3.2 Neoclassical Optimal Growth 6
1.4 Regional Allocation of Public Investment 7
1.4.1 The Dynamics of Capital Formation 8
1.4.2 Population Dynamics 10
1.4.3 Technological Change 12
1.4.4 Criterion Functional and Final Form of the Model 13
1.5 Dynamic Telecommunications Flow Routing 14
1.5.1 Assumptions and Notation 14
1.5.2 Flow Propagation Mechanism 15
1.5.3 Path Delay Operators 17
1.5.4 Dynamic System Optimal Flows 18
1.5.5 Additional Constraints 18
1.5.6 Final Form of the Model 20
1.6 Brief History of Dynamic Games 20
1.7 Dynamic User Equilibrium for Vehicular Networks 21
1.8 Dynamic Oligopolistic Network Competition 22
1.8.1 Notation 23
1.8.2 Extremal Problems and the Nash Game 23
1.8.3 Differential Variational Inequality Formulation 26
1.9 Revenue Management and Nonlinear Pricing 27
1.9.1 The Decision Environment 27
1.9.2 The Role of Denial-of-Service Costs and Refunds 29
1.9.3 Firms' Extremal Problem 29
1.10 The Material Ahead 30
List of References Cited and Additional Reading 31

2 Nonlinear Programming and Discrete-Time Optimal Control 33
2.1 Nonlinear Program Defined 34
2.2 Other Types of Mathematical Programs 35
2.3 Necessary Conditions for an Unconstrained Minimum 37
2.4 Necessary Conditions for a Constrained Minimum 38
2.4.1 The Fritz John Conditions 38
2.4.2 Geometry of the Kuhn-Tucker Conditions 39
2.4.3 The Lagrange Multiplier Rule 41
2.4.4 Motivating the Kuhn-Tucker Conditions 43
2.5 Formal Derivation of the Kuhn-Tucker Conditions 46
2.5.1 Cones and Optimality 47
2.5.2 Theorems of the Alternative 48
2.5.3 The Fritz John Conditions Again 49
2.5.4 The Kuhn-Tucker Conditions Again 50
2.6 Sufficiency, Convexity, and Uniqueness 52
2.6.1 Quadratic Forms 52
2.6.2 Concave and Convex Functions 53
2.6.3 Kuhn-Tucker Conditions Sufficient 58
2.7 Generalized Convexity and Sufficiency 60
2.8 Numerical and Graphical Examples 62
2.8.1 LP Graphical Solution 62
2.8.2 NLP Graphical Example 64
2.8.3 Nonconvex, Nongraphical Example 66
2.8.4 A Convex, Nongraphical Example 68
2.9 Discrete-Time Optimal Control 69
2.9.1 Necessary Conditions 71
2.9.2 The Minimum Principle 74
2.9.3 Discrete Optimal Control Example 75
2.10 Exercises 77
List of References Cited and Additional Reading 78

3 Foundations of the Calculus of Variations and Optimal Control 79
3.1 The Calculus of Varations 80
3.1.1 The Space $C^1\left[t_0, t_f\right]$ 80
3.1.2 The Concept of a Variation 80
3.1.3 Fundamental Lemma of the Calculus of Variations 82
3.1.4 Derivation of the Euler-Lagrange Equation 85
3.1.5 Additional Necessary Conditions in the Calculus of Variations 87
3.1.6 Sufficiency in the Calculus of Variations 93
3.1.7 Free Endpoint Conditions in the Calculus of Variations 95
3.1.8 Isoperimetric Problems in the Calculus of Variations 95
3.1.9 The Beltrami Identity for $\frac{\partial f_0}{\partial t} = 0$ 96

3.2 Calculus of Variations Examples 97
3.2.1 Example of Fixed Endpoints in the Calculus of Variations . 98
3.2.2 Example of Free Endpoints in the Calculus of Variations .. 99
3.2.3 The Brachistochrone Problem 100
3.3 Continuous-Time Optimal Control 103
3.3.1 Necessary Conditions for Continuous-Time Optimal Control 105
3.3.2 Necessary Conditions with Fixed Terminal Time, No Terminal Cost, and No Terminal Constraints 109
3.3.3 Necessary Conditions When the Terminal Time Is Free.....111
3.3.4 Necessary Conditions for Problems with Interior Point Constraints 113
3.3.5 Dynamic Programming and Optimal Control 114
3.3.6 Second-Order Variations in Optimal Control 117
3.3.7 Singular Controls 119
3.3.8 Sufficiency in Optimal Control 120
3.4 Optimal Control Examples 124
3.4.1 Simple Example of the Minimum Principle 124
3.4.2 An Example Involving Singular Controls 127
3.4.3 Approximate Solution of Optimal Control Problems by Time Discretization 130
3.4.4 A Two-Point Boundary-Value Problem 130
3.4.5 Example with Free Terminal Time 134
3.5 The Linear-Quadratic Optimal Control Problem 138
3.5.1 LQP Optimality Conditions 138
3.5.2 The HJPDE and Separation of Variables for the LQP 140
3.5.3 LQP Numerical Example.....141
3.5.4 Another LQP Example 142
3.6 Exercises.....144
List of References Cited and Additional Reading 145

4 Infinite Dimensional Mathematical Programming 147
4.1 Elements of Functional Analysis 148
4.1.1 Notation and Elementary Concepts 148
4.1.2 Topological Vector Spaces 149
4.1.3 Convexity 157
4.1.4 The Hahn-Banach Theorem 158
4.1.5 Gâteaux Derivatives and the Gradient of a Functional.....159
4.1.6 The Fréchet Derivative 162
4.2 Variational Inequalities and Constrained Optimization of Functionals 163
4.3 Continuous-Time Optimal Control 165
4.3.1 Analysis Based on the G-Derivative 166
4.3.2 Variational Inequalities as Necessary Conditions 169

4.4 Optimal Control with Time Shifts 174
4.4.1 Some Preliminaries 175
4.4.2 The Optimal Control Problem of Interest 176
4.4.3 Change of Variable 176
4.4.4 Necessary Conditions for Time-Shifted Problems 177
4.4.5 A Simple Abstract Example....183
4.5 Derivation of the Euler-Lagrange Equation 185
4.6 Kuhn-Tucker Conditions for Hilbert Spaces 186
4.7 Mathematical Programming Algorithms 190
4.7.1 The Steepest Descent Algorithm....190
4.7.2 The Projected Gradient Algorithm....198
4.7.3 Penalty Function Methods 204
4.7.4 Example of the Steepest Descent Algorithm 206
4.7.5 Example of the Gradient Projection Algorithm 209
4.7.6 Penalty Function Example 214
4.8 Exercises....216
List of References Cited and Additional Reading 217

5 Finite Dimensional Variational Inequalities and Nash Equilibria 219
5.1 Some Basic Notions....220
5.2 Nash Equilibria and Normal Form Games 220
5.3 Some Related Nonextremal Problems....222
5.3.1 Nonextremal Problems and Programs 223
5.3.2 Kuhn-Tucker Conditions for Variational Inequalities....224
5.3.3 Variational Inequality and Complementarity Problem Generalizations 226
5.3.4 Relationships Among Nonextremal Problems 226
5.3.5 Variational Inequality Representation of Nash Equilibrium 231
5.3.6 User Equilibrium 231
5.3.7 Existence and Uniqueness....235
5.4 Sensitivity Analysis of Variational Inequalities....237
5.5 The Diagonalization Algorithm 239
5.5.1 The Algorithm 241
5.5.2 Converence of Diagonalization 242
5.5.3 A Nonnetwork Example of Diagonalization 243
5.6 Gap Function Methods for $VI\,(F, \Lambda)$ 248
5.6.1 Gap Function Defined 248
5.6.2 The Auslender Gap Function 249
5.6.3 Fukushima-Auchmuty Gap Functions....250
5.6.4 The D-Gap Function....251
5.6.5 Gap Function Numerical Example....253
5.7 Other Algorithms for $VI\,(F, \Lambda)$ 255
5.7.1 Methods Based on Differential Equations....256
5.7.2 Fixed-Point Methods 257

5.7.3 Generalized Linear Methods 258
5.7.4 Successive Linearization with Lemke's Method 259
5.8 Computing Network User Equilibria 260
5.9 Exercises 263
List of References Cited and Additional Reading 263

6 Differential Variational Inequalities and Differential Nash Games 267
6.1 Infinite-Dimensional Variational Inequalities 268
6.2 Differential Variational Inequalities 271
6.2.1 Problem Definition 271
6.2.2 Naming Conventions 272
6.2.3 Regularity Conditions for *DVI(F, f, Ψ, U, x_0, t_0, t_f)* 273
6.2.4 Necessary Conditions 274
6.2.5 Existence 276
6.2.6 Nonlinear Complementarity Reformulation 276
6.3 Differential Nash Games 277
6.3.1 Differential Nash Equilibrium 277
6.3.2 Generalized Differential Nash Equilibrium 281
6.4 Fixed-Point Algorithm 282
6.4.1 Formulation 282
6.4.2 The Unembellished Algorithm 283
6.4.3 Solving the SubProblems 286
6.4.4 Numerical Example 287
6.5 Descent in Hilbert Space with Gap Functions 289
6.5.1 Gap Functions in Hilbert Spaces 289
6.5.2 D-gap Equivalent Optimal Control Problem 292
6.5.3 Numerical Example 297
6.6 Differential Variational Inequalities with Time Shifts 298
6.6.1 Necessary Conditions 300
6.6.2 Fixed-Point Formulation and Algorithm 303
6.6.3 Time-Shifted Numerical Examples 305
6.7 Exercises 310
List of References Cited and Additional Reading 311

7 Optimal Economic Growth 313
7.1 Alternative Models of Optimal Economic Growth 314
7.1.1 Ramsey's 1928 Model 314
7.1.2 Optimal Growth with the Harrod-Domar Model 315
7.1.3 Neoclassical Optimal Growth 316
7.2 Optimal Regional Growth Based on the Harrod-Domar Model 318
7.2.1 Tax Rate as the Control 325
7.2.2 Tax Rate and Public Investment as Controls 328
7.2.3 Equal Public and Private Savings Ratios 333
7.2.4 Sufficiency 339

7.3 A Computable Theory of Regional Public Investment Allocation ...343
7.3.1 The Dynamics of Capital Formation ...344
7.3.2 Population Dynamics ...345
7.3.3 Technological Change ...347
7.3.4 Criterion Functional and Final Form of the Model ...348
7.3.5 Numerical Example Solved by Time Discretization ...349
7.4 Exercises ...350
List of References Cited and Additional Reading ...351

8 Production Planning, Oligopoly and Supply Chains ...353
8.1 The Aspatial Price-Taking Firm ...354
8.1.1 Optimal Control Problem for Aspatial Perfect Competition ...355
8.1.2 Numerical Example of Aspatial Perfect Competition ...355
8.1.3 The Aspatial Price Taking Firm with a Terminal Constraint on Inventory ...358
8.2 The Aspatial Monopolistic Firm ...362
8.2.1 Necessary Conditions for the Aspatial Monopoly ...363
8.2.2 Numerical Example ...364
8.3 The Monopolistic Firm in a Network Economy ...367
8.3.1 The Network Firm's Extremal Problem ...367
8.3.2 Discrete-Time Approximation ...370
8.3.3 Numerical Example ...371
8.3.4 Solution by Discrete-Time Approximation ...373
8.3.5 Solution by Continuous-Time Gradient Projection ...373
8.4 Dynamic Oligopolistic Spatial Competition ...376
8.4.1 Some Background and Notation ...377
8.4.2 The Firm's Objective and Constraints ...378
8.4.3 The DVI Formulation ...380
8.4.4 Discrete-Time Approximation ...384
8.4.5 A Comment About Path Variables ...386
8.4.6 Numerical Example ...386
8.4.7 Interpretation of Numerical Results ...389
8.5 Competitive Supply Chains ...395
8.5.1 Inverse Demands ...395
8.5.2 Producers' Extremal Problem ...396
8.5.3 Retailers' Extremal Problem ...399
8.5.4 Supply Chain Extremal Problem ...400
8.5.5 The Differential Variational Inequality ...401
8.5.6 The DVI ...404
8.5.7 Numerical Example ...405
8.6 Exercises ...408
List of References Cited and Additional Reading ...409

9 Dynamic User Equilibrium 411
9.1 Some Background 412
9.2 Arc Dynamics 413
9.2.1 Dynamics Based on Arc Exit Flow Functions 413
9.2.2 Dynamics with Controlled Entrance and Exit Flows 414
9.2.3 Cell Transmission Dynamics 415
9.2.4 Dynamics Based on Arc Exit Time Functions 416
9.2.5 Constrained Dynamics Based on Proper Flow Propagation Constraints 418
9.3 The Measure-Theoretic Nature of DUE 421
9.4 The Infinite-Dimensional Variational Inequality Formulation 423
9.5 When Delays Are Exogenous 428
9.6 When the Delay Operators Are Endogenous 435
9.6.1 Nested Operators 437
9.6.2 The Problem Setting 438
9.6.3 Analysis 439
9.6.4 Computation with Endogenous Delay Operators 445
9.7 Conclusions 452
9.8 Exercises 453
List of References Cited and Additional Reading 455

10 Dynamic Pricing and Revenue Management 457
10.1 Dynamic Pricing with Fixed Inventories 458
10.1.1 Infinite-Dimensional Variational Inequality Formulation 460
10.1.2 Restatement of the Isoperimetric Constraints 463
10.1.3 Differential Variational Inequality Formulation 464
10.1.4 Numerical Example 464
10.2 Revenue Management as an Evolutionary Game 467
10.2.1 Assumptions and Notation 468
10.2.2 Demand Dynamics 469
10.2.3 Constraints 470
10.2.4 The Firm's Optimal Control Problem 471
10.2.5 Differential Quasivariational Inequality Formulation 473
10.2.6 Numerical Example 475
10.3 Network Revenue Management 481
10.3.1 Discrete-Time Notation 481
10.3.2 Demand Functions 483
10.3.3 Denial-of-Service Costs and Refunds 485
10.3.4 Firms' Extremal Problem 485
10.3.5 Market Equilibrium Problem as a Quasivariational Inequality 487
10.3.6 Numerical Example 488
10.4 Exercises 490
List of References Cited and Additional Reading 492

Index 495

Preface

This is a book about continuous time optimal control and its extension to a certain class of dynamic games known as open loop differential Nash games. Its intended audience is students and researchers wishing to model and compute in continuous time. The presentation is meant to be accessible to a wide audience. Accordingly, the presentation does not always rest on the most general and least restrictive regularity assumptions.

This book may be used by those with little mathematical preparation beyond introductory differential and integral calculus and a first course in ordinary differential equations. Nonetheless, prior exposure to nonlinear programming is desirable. For those without that prior exposure, a chapter that reviews the foundations of NLP is included.

The exercises at the end of each chapter should be attempted by anyone seeking mastery of the material emphasized in this book. The exercises are in some cases very challenging, yet they accurately represent the kinds of problems one faces in building and applying dynamic models based on optimal control theory and dynamic non-cooperative game theory.

Chapter 1 provides some insight into the history of dynamic optimization and differential games, as well as a preview of the applications that are covered in the book. Chapter 2 provides a review of finite dimensional nonlinear programming, while Chapters 3 and 4 present the foundations of continuous time optimal control and infinite dimensional mathematical programming. Chapter 5 provides a condensed treatment of finite dimensional Nash games and their representation as variational inequalities. Chapter 6 presents the foundations of open loop dynamic Nash games and their representation as differential variational inequalities. Chapters 7, 8, 9 and 10 are devoted, respectively, to the following applications: economic growth theory; production planning and supply chains; dynamic user equilibrium; and pricing and revenue management.

I regret that the press of time has not permitted me to cover differential Stackelberg games or to include chapters on stochastic differential games. It is my hope that these and other incomplete aspects of the book may someday be overcome in a revised edition.

The preparation of this book as been made possible by several grants from the National Science Foundation and the continuous support of the Harold and Inge

Marcus Chaired Professorship. I have also been aided by several generations of graduate students and graduate research assistants at Penn State, George Mason University, the University of Pennsylvania and MIT. In this regard, I am especially indebted to Changhyun Kwon, Pat Harker, David Bernstein, Enrique Fernandez, H. J. Cho, C. C. Lin, Niko Kydes, Reeto Mookherjee, Taeil Kim, B.W. Wie and Ilsoo Lee for their assistance with countless theoretical, numerical and applied explorations of optimal control and game theory. I am also deeply grateful for the aid and support of my wife, Joyce, and the companionship of Dakota, Coaly Bay, Montana, and Nevada.

University Park, Pennsylvania *Terry L. Friesz*
2010

Chapter 1
Introduction

In this book we present the theory of continuous-time dynamic optimization, covering the classical calculus of variations, the modern theory of optimal control, and their linkage to infinite-dimensional mathematical programming. We present an overview of the main classes of practical algorithms for solving dynamic optimization problems and develop some facility with the art of formulating dynamic optimization models. Upon completing our study of dynamic optimization, we turn to dynamic Nash games. Our coverage of dynamic games emphasizes continuous-time variational inequalities and subsumes portions of the classical theory of differential games.

This book, although it may be used as a text, is meant to be a reference and guide to engineers, applied mathematicians, and social scientists whose work necessitates a background in dynamic optimization and differential games. It provides a detailed exposition of several original dynamic and game-theoretic mathematical models of operations research and management science applications of current importance, including revenue management, supply chain management, and dynamic traffic assignment.

Ideally, the reader should have a prior mastery of necessary and sufficient conditions of finite dimensional (nonlinear) mathematical programming and of the main algorithmic philosophies used to solve nonlinear programs. The reasons for needing this background in nonlinear programming will become abundantly clear as we proceed. Nonetheless, many individuals without significant prior exposure to nonlinear programming will likely wish to become familiar with and apply the theory of dynamic optimization and/or the theory of differential games exposited in this book. Consequently, in Chapter 2, we review nonlinear programming to a depth that will allow the reader without a prior background in nonlinear programming to understand a significant fraction of the material in subsequent chapters of this book. Such readers must keep in mind, as we re-emphasize in Chapter 2, that full preparation in nonlinear programming cannot be obtained from Chapter 2 alone. In fact, for full comprehension of the material contained in subsequent chapters, one needs to have successfully completed a formal course in nonlinear programming at the graduate level. If one's background is limited to a one-time exposure to the material of Chapter 2 certain analyses and results presented in subsequent chapters will be difficult to understand and in some instances will be opaque.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_1, © Springer Science+Business Media, LLC 2010

The following is an outline of the principal topics covered in this chapter:

Section 1.1: Brief History of the Calculus of Variations and Optimal Control. We give a brief summary of several hundred years of mathematical research, emphasizing the origins of the calculus of variations and the special insights offered by its modern generalization, the theory of optimal control.

Section 1.2: The Brachistochrone Problem. We review a specific version of the class of minimum time problems that launched the calculus of variations in the 17th century.

Section 1.3: Optimal Economic Growth. We study the paradigm of neoclassical optimal economic growth theory, as well as its historical antecedent, Ramsey's 1928 model.

Section 1.4: Regional Allocation of Public Investment. We show how neoclassical optimal economic growth theory may be extended to consider interacting regions of a national economy.

Section 1.5: Dynamic Telecommunications Flow Routing. We present a detailed deterministic model of message flow routing in a dynamic telecommunications network.

Section 1.6: Brief History of Differential Games. We give a brief overview of the history of mathematical games with explicit dynamics, emphasizing the contribution of Issacs and perspectives based on differential variational inequalities.

Section 1.7: Dynamic User Equilibrium for Vehicular Networks. We show how one of the most challenging applied problems of present-day operations research, the so-called dynamic user equilibrium problem for predicting urban traffic flows, may be viewed as a differential Nash game and expressed as a differential variational inequality.

Section 1.8: Dynamic Oligopolistic Network Competition. We show how static notions of oligopolistic competition may be generalized to describe a Nash-type equilibrium on a network whose flows are driven by inventory and shipping considerations.

Section 1.9: Revenue Management and Nonlinear Pricing. We give one example of pricing in a dynamic, nonlinear, competitive setting. Again the model takes the form of a differential variational inequality.

1.1 Brief History of the Calculus of Variations and Optimal Control

Most of the abstract mathematical problems and dynamic optimization models we will study belong to the broad topic of inquiry known as the *calculus of variations*. The calculus of variations, to give a very crude definition that we will refine later, is a

formalism for analyzing and solving extremal problems in function spaces. As such the unknowns we are seeking are themselves functions and the extremal criterion is a *functional* – that is, a function of other functions. The independent variable is a scalar that, in applications, typically corresponds to time, although this scalar may in fact be any independent variable that is useful for a parametric representation of the dependent variables of interest.

Another notable feature of the calculus of variations is that the extremal criterion is frequently an integral. This means that many continuous-time optimization models that involve present value calculations may be viewed as calculus of variations problems. The limits of integration of the extremal integral may be fixed or unknowns to be determined. This allows models to be created that determine not only the trajectory through time of an optimal process but also the time of its initiation as well as the time of its termination. It is widely agreed that the calculus of variations was born almost immediately after the creation of the Newton-Leibnitz calculus familiar to almost all first-year college students. In fact, Newton and Leibnitz played a role in the early development of the calculus of variations, although the Bernoulli brothers, Euler, and Lagrange are generally credited with the innovations that we loosely refer to as the classical calculus of variations. In the classical calculus of variations, side conditions, or constraints – as a person with modern training in mathematical programming would call them – can be accommodated but only with some difficulty. Perhaps the side conditions that pose the greatest challenge for the classical theory are functional equations, especially differential equations. By contrast, optimal control theory – which may be thought of as a modern calculus of variations – is able to treat differential equation constraints with relative ease. Optimal control theory also employs much more hygienic notational conventions and distinguishes between two classes of decision variables: control variables and state variables. This distinction meshes perfectly in many economic and techno-economic applications with intuitive notions of what variables may be used to set policy (control variables) and what variables describe the implications of policy (state variables).

Optimal control theory in effect originates with the publication of *The Mathematical Theory of Optimal Processes* by Pontryagin, Boltyanskii, Gamkrelidze, and Mishchenko in 1958 in Russian and 1962 in English. This book not only provided a powerful formalism for including all types of constraints in calculus of variations problems but also introduced the important theoretical result known as the *maximum principle* (or *minimum principle* when minimizing). The maximum principle allows the decomposition of a dynamic optimization problem in the form of an optimal control problem into a set of static problems, one for each instant of time. The optimal solution of the instantaneous subproblems can be shown to give the optimal solution of the original dynamic problem. The power of this result cannot be overstated – it is one of the most important intellectual achievements of the twentieth century and one of the most important mathematical discoveries of all time.

To satisfy diverse applications interests, we will provide in the present chapter some example dynamic optimization models that are historically important and/or of current topical interest. The models we consider are the following:

1. the brachistochrone problem;
2. optimal economic growth;
3. regional investment allocation; and
4. dynamic telecommunications flow routing.

When studying these models, our perspective will be exclusively deterministic in both this and subsequent chapters.

1.2 The Brachistochrone Problem

The archetypal calculus of variations problem is the so-called *brachistochrone*[1] *problem*:

> Imagine a bead slides along a frictionless wire guideway connecting two fixed points A and B in a constant gravitational field. The speed of the bead along the wire is V_0 at point A. What shape must the wire have to produce a minimum-time path between A and B?

History records this as one of the first, if not the first, calculus of variations problem; a version of the problem was posed by Johann Bernoulli in 1696 and explored further by his brother Jakob Bernoulli, as well as by Newton, L'Hôpital and Euler. The solution is a cycloid, lying in the vertical plane and passing through A and B.

Let $y(x)$ denote the vertical position of the bead as a function of its horizontal position x. Working through the equations of motion for this problem, it can be shown that the speed of the bead at depth y below the origin is $(2gy)^{\frac{1}{2}}$ and that the time of passage is the functional

$$J\left(x, y, y'\right) = \frac{1}{(2g)^{\frac{1}{2}}} \int_0^x \left[1 + \left(\frac{dy(\xi)}{dt}\right)^2\right] [y(\xi)]^{-\frac{1}{2}} \, d\xi$$

where ξ is a dummy variable of integration and g is the acceleration due to gravity. It is of course the minimization

$$\min J\left(x, y, y'\right)$$

that constitutes the brachistochrone problem. If one studies the historical writings on the brachistochrone problem, it becomes apparent that there are many variants and specializations of the basic version we have presented. For this reason, some of the older books refer, rightly so, to the *brachistochrone problems*.

[1] The word *brachistochrone* is Greek and means "shortest time."

1.3 Optimal Economic Growth

We consider two models of optimal economic growth that are interesting in their own right and that provide an introduction to the type of thinking characteristic of dynamic continuous-time optimization models.

1.3.1 Ramsey's 1928 Model

In a very famous paper Ramsey (1928) proposed the idea of a *bliss point*, an accumulation point of a sequence of consumption decisions representing the nonattainable, ideal consumption goal of the consumer. The bliss point, $B > 0$, has the same units as utility and obeys

$$B = \sup [U(c) : c \geq 0]$$

where c is the consumption level of a representative member of society and $U(c)$ is the utility experienced as a result of that consumption. Because there are N identical members of society, maximization of social welfare is assured by

$$\max J = \int_0^\infty [U(c) - B]\, dt \tag{1.1}$$

The relevant dynamics are obtained from a neoclassical production function

$$Y = F(K, L) \tag{1.2}$$

where Y is output and K and L are time-varying capital input and constant labor input, respectively. The neoclassical nature of (1.2) means that $F(.,.)$ is homogeneous of degree one; that is

$$F(\alpha K, \alpha L) = \alpha F(K, L)$$

for α a positive scalar. Taking labor (population) to be fixed, per capita output may be expressed as

$$\begin{aligned} y = \frac{Y}{L} &= \frac{1}{L} F(K, L) \\ &= F\left(\frac{K}{L}, 1\right) = f(k) \end{aligned}$$

where

$$k \equiv \frac{K}{L} \quad \text{and} \quad f(k) \equiv F\left(\frac{K}{L}, 1\right)$$

Thus, we have

$$\begin{aligned}\frac{dk}{dt} &= \frac{d}{dt}\left(\frac{K}{L}\right) = \frac{1}{L}\frac{dK}{dt} \\ &= \frac{F(K,L)}{L} - \frac{C}{L} - \delta\frac{K}{L} \\ &= f(k) - c - \delta k \end{aligned} \tag{1.3}$$

where I is investment, C is total consumption, c is per capita consumption, and δ is the rate of depreciation of capital. Obviously, k is per capita capital. Note that per capita output $f(k)$ is expressed in terms of the single state variable k, a fact made possible by the homogeneous-of-degree-one property of the production function. In deriving (1.3), we make use of well-known macroeconomic identities relating the rate of change of capital stocks dK/dt, investment I, consumption C, and capital depreciation δK:

$$\begin{aligned}\frac{dK}{dt} &= I - \delta K \\ &= Y - C - \delta K\end{aligned}$$

The above development allows us to state Ramsey's optimal growth model in the following form:

$$\max J = \int_0^\infty [U(c) - B]\, dt \tag{1.4}$$

subject to

$$\frac{dk}{dt} = f(k) - c - \delta k \tag{1.5}$$

$$k(0) = k_0 \tag{1.6}$$

The model (1.4), (1.5), and (1.6) is an optimal control problem with state variable k and control variable c.

1.3.2 Neoclassical Optimal Growth

Ramsey's work can be criticized from the points of view that population is not time varying, that there is no discounting, and that the concept of a bliss point contradicts the nonsatiation axiom of utility theory. These criticisms are quite easy to remedy and lead to the so-called neoclassical model of optimal growth. One of the clearest and most succinct expositions of the neoclassical theory of optimal economic growth is contained in the book by Arrow and Kurz (1970).

To express the neoclassical model of optimal growth, we postulate that population has a constant proportionate growth rate π obeying

$$\frac{1}{L}\frac{dL}{dt} = \pi \Longrightarrow L = L_0 \exp(-\pi t) \tag{1.7}$$

where $L_0 = L(0)$ is the initial population (labor force). We then construct dynamics for *per capita* capital $k = K/L$ in a fashion highly similar to that employed for Ramsey's model. In particular we write

$$\begin{aligned}\frac{dk}{dt} &= \frac{d}{dt}\left(\frac{K}{L}\right) = \frac{1}{L}\frac{dK}{dt} - \frac{K}{L}\frac{1}{L}\frac{dL}{dt} \\ &= \frac{F(K,L)}{L} - \frac{C}{L} - \delta\frac{K}{L} - \pi\frac{K}{L} \\ &= f(k) - c - \delta k - \pi k\end{aligned}$$

where now

$$f(k) = F\left(\frac{K}{L}, 1\right)$$

Consequently, the neoclassical optimal growth model is

$$\max J = \int_0^\infty \exp(-\rho t)\, U(c)\, dt \tag{1.8}$$

subject to

$$\frac{dk}{dt} = f(k) - c - \delta k - \pi k \tag{1.9}$$

$$k(0) = k_0 \tag{1.10}$$

where ρ is the constant nominal rate of discount. The model comprised of (1.8), (1.9), and (1.10) is again an optimal control problem. The neoclassical optimal growth model is itself open to criticism, especially as regards its assumption of a constant returns to scale technology.

1.4 Regional Allocation of Public Investment

As another example, let us next consider a model for optimal regional allocation of public investment, derived from taxes on private sector earnings. That model in its most general form includes the following characteristics:

1. growth dynamics involve no constant returns assumption and allow increasing returns;
2. there is an equilibrium in all capital markets;

3. private and public capital, the latter allowing infrastructure investment decisions to be modeled, are distinguished from each other;
4. population evolves over time in accordance with a Hotelling-type diffusion model that includes births, deaths, and location-specific ecological carrying capacities;
5. capital augmenting technological change is allowed and is endogenous in nature;
6. regulatory and fiscal policy constraints may be imposed; and
7. the optimization criterion is the present value of the national income time stream.

The model we propose is partly based on the spatial disaggregation of macroeconomic identities relating the rate of change of capital stocks to investments and depreciation. This perspective on disaggregation to create coupled differential equations describing regional growth can be traced back to Datta-Chaudhuri (1967), Sakashita (1967), Ohtsuki (1971), Domazlicky (1977), Bagchi (1984), and Friesz and Luque (1987), although the assumptions we make regarding production functions and technological change are extensions of this antecedent literature. The below model also differs from historical regional growth models in that we do not rely on the assumption of a constant proportionate rate of labor force growth for each region. The constant proportionate growth (CPG) model of labor and population allows the dynamics for population to be uncoupled from those of capital formation and technological change. As a consequence, in CPG models population always grows exponentially with respect to time and shows no response to changes in population density, capital or regional income. As an alternative, we replace the unrealistic CPG model of population growth with a Hotelling-type model that includes the effects of spatial diffusion and of ecological carrying capacities of individual regions and is intrinsically coupled to the dynamics of capital formation.

1.4.1 The Dynamics of Capital Formation

Basic macroeconomic identities can be used to describe the relationship of the rate of change of capital to investment, output and savings. In the simplest framework, output is a function of capital and labor, and the rate of change of capital is equated to investment less any depreciation of capital that may occur. That is:

$$\frac{dK}{dt} = I - \delta K \tag{1.11}$$

where K is capital, dK/dt is the time rate of change of capital, I is investment, and δ is an abstract depreciation rate. Subsequently δ_p will be the depreciation rate of private capital and δ_g the depreciation rate of public capital. Of course (1.11) is an aspatial model. It is also important to recognize that (1.11) is an equilibrium

model for which the supply of capital is exactly balanced against the demand for capital. We shall maintain this assumption of capital market equilibrium throughout the development that follows. It is also worth noting that (1.11) is the foundation of the much respected work by Arrow and Kurz (1970) exploring the interdependence of aspatial private and public sector growth dynamics. To spatialize (1.11) as well as to introduce a distinction between the public and private sectors we write:

$$\frac{dK_i^p}{dt} = I_i^p - \delta_p K_i^p \quad \text{and} \quad \frac{dK_i^g}{dt} = I_i^g - \delta_g K_i^g \tag{1.12}$$

where the subscript $i \in [1, N]$ refers to the ith of N regions, and the superscripts p and g refer to the private and public (governmental) sectors, respectively.

Further detail can be introduced into the above dynamics by defining c_i to be the consumption rate of region i and r to be a tax rate imposed by the central government on each region's output. We also define η_i to be the share of tax revenues allocated to subsidize private investments in region i, and ν_i to be the share of tax revenues allocated to public (infrastructure) investments in region i. Also, Y_i will be the output of region i. To keep the presentation simple, we assume that all capital (private as well as public) is immobile, although this assumption can be relaxed at the expense of more complicated notation. Consequently, the following two identities hold:

$$I_i^p = (1 - c_i - r)\, Y_i + \eta_i r \sum_{j=1}^{N} Y_j \quad \text{and} \quad I_i^g = \nu_i r \sum_{j=1}^{N} Y_j \tag{1.13}$$

By virtue of the definitions of η_i and ν_i, we have the following constraints:

$$\sum_{i=1}^{N} (\eta_i + \nu_i) = 1 \tag{1.14}$$

and

$$0 \leq \eta_i \leq 1 \quad \text{and} \quad 0 \leq \nu_i \leq 1 \tag{1.15}$$

which require that allocations cannot exceed the tax revenues collected and must be nonnegative.[2] We further assume that the i^{th} region's intrinsic technology, ignoring for the moment technological innovation, is described by a production function of the form

$$Y_i = F_i \left(K_i^p, K_i^g, L_i\right) \tag{1.16}$$

where L_i is the labor force (population) of the i^{th} region.

[2] Other tax schemes, such as own-region taxes can easily be described. The one chosen here is meant to be illustrative.

It follows at once from (1.12) and (1.16) that the dynamics for the evolution of private and public sector capital are, respectively:

$$\frac{dK_i^p}{dt} = (1 - c_i - r)\, F_i\left(K_i^p, K_i^g, L_i\right) + \eta_i r \sum_{j=1}^{N} F_j\left(K_j^p, K_j^g, L_j\right) - \delta_p K_i^p \qquad \forall i \in [1, N] \tag{1.17}$$

$$\frac{dK_i^g}{dt} = \nu_i r \sum_{j=1}^{N} F_j\left(K_j^p, K_j^g, L_j\right) - \delta_g K_i^g \qquad \forall i \in [1, N] \tag{1.18}$$

It is important to note that we have made no assumption regarding constant returns to scale in articulating the above dynamics.

1.4.2 Population Dynamics

Traditionally, the literature on neoclassical economic growth, as we noted above, has assumed a constant proportionate rate of labor force growth. As population and labor force are typically treated as synonymous, this means that models from this literature employ quite simple population growth models of the form

$$\frac{dL_i}{dt} = \pi_i L_i \qquad L_i(0) = L_i^0 \tag{1.19}$$

for every region $i \in N$ where π_i is a constant. This means that population (labor force) always grows according to the exponential law $L_i(t) = L_i^0 e^{\pi_i t}$ regardless of any other assumptions employed. The CPG assumption is decidedly unrealistic, limits the policy usefulness of economic growth models based on it, and calls out to be replaced with a richer model of population and labor force change over time and space.

We replace (1.19) with a spatial diffusion model of the Hotelling type[3]. In Hotelling-type population models, migration is based on the noneconomic notion of diffusion wherein populations seek spatial niches that have been previously unoccupied. This means that, unlike (1.19), population will not become inexorably denser at a given point in space, but rather that population density may rise and fall over time. Yet because we will link this diffusion process to the capital formation process, there will be a potential for population to concentrate where infrastructure agglomeration economies occur. Furthermore, we will employ a version of the spatial diffusion process that includes a logistic model of birth/death processes

[3] See Hotelling (1978).

and specifically incorporates the ecological carrying capacity of each location alternative. These features will inform and be informed by the capital dynamics (1.17) and (1.18), resulting in an economic growth model that is intrinsically more realistic than would result from rote adherence to the neoclassical paradigm.

Hotelling's original model is in the form of a partial differential equation which is very difficult to solve for realistic spatial boundary conditions and is not readily coupled with ordinary differential equations such as (1.17) and (1.18). In Puu (1989) and Puu (1997), a multiregion alternative to Hotelling's model is suggested; that alternative captures key features of the diffusion process and the birth/death process in a more tractable mathematical framework. Specifically, Puu (1989) proposes, if the population of region i is denoted as P_i, the following dynamics:

$$\frac{dP_i}{dt} = \gamma_i P_i \left(\zeta_i - P_i\right) + \sum_{j \neq i} \lambda_j \left(P_j - P_i\right) \qquad \forall i \in [1, N] \tag{1.20}$$

where γ_i, ζ_i, and λ_i are positive exogenous parameters to him. The idea here is that the term $\sum_{j \neq i} \lambda_j \left(P_j - P_i\right)$ is roughly analogous to diffusion in that it draws population from regions with higher population density toward regions with lower population density. Typically λ_j is referred to as the *coefficient of diffusion* for region $j \in N$. The entity ζ_i is sometimes called the *fitness measure* and describes the *ecological carrying capacity* of region $i \in N$; its units are population. The parameter γ_i ensures dimensional consistency and has the units of $(time)^{-1}$. Clearly this model is not equivalent to Hotelling's, but it does capture the essential ideas behind diffusion-based population growth and migration and is substantially more tractable from a computational point of view since (1.20) is a system of ordinary (as opposed to partial) differential equations.

Moreover, the population dynamics (1.20) can be considerably enriched by allowing the fitness measure to be locationally and infrastructurally specific, as we now show. Specifically, we postulate that

$$\zeta_i = V_i \left(K_i^g, t\right) + \Psi_i \left(t\right) \tag{1.21}$$

where $V_i \left(K_i^g, t\right)$ describes the effect of infrastructure on carrying capacity and $\Psi_i \left(t\right)$ is the natural or ambient carrying capacity that exists in the absence of infrastructure investment. It is important to understand that by "carrying capacity" we mean the population that a region can sustain. As such, (1.21) expresses the often-made observation that each individual region is naturally prepared to support a specific population level, and that level may vary with time and be conditioned by manmade infrastructure. Puu (1989) observes that population models such as that presented above have one notable shortcoming: it is possible, for certain initial conditions, that population trajectories will include periods of negative population. Negative population is, of course, meaningless, and population trajectories with this property cannot be accepted as realistic. Consequently, we must include in the final optimal control formulation a state space constraint that forces population to remain nonnegative.

1.4.3 Technological Change

None of the above presentation depends on the idea of balanced growth or the assumption that a long-run equilibrium exists in the usual sense. Neither do we assume that technological progress must obey some type of neutrality. When introducing technological progress, we have a free hand to explore any type of technological progress. In particular, we are not restricted to labor augmenting progress or progress that enhances overall output, and we can explore capital augmenting progress, a natural choice since our interest is in the role of public capital (infrastructure). In fact, we shall concentrate on technological progress that augments K_i^g. Consequently for each region i, we shall update the associated production function by making the substitution

$$K_i^g \Longrightarrow \Phi(t) K_i^g \tag{1.22}$$

where Φ is a scalar function of time describing the extent of infrastructure augmenting technological progress. Technological progress for our model is endogenously generated through separate dynamics for Φ. We postulate that public capital augmenting technological progress occurs when the ratio of national output to public-capital falls below some threshold and is zero when the ratio exceeds that threshold. That is, the rate of technological progress $d\Phi/dt$ obeys

$$\frac{d\Phi}{dt} > 0 \quad \text{if} \quad \sigma \leq \Theta \tag{1.23}$$

$$\frac{d\Phi}{dt} = 0 \quad \text{if} \quad \sigma > \Theta \tag{1.24}$$

where σ is the output/public-capital ratio and $\Theta \in \Re_+^1$ is a known reference threshold which determines the need for and the fact of technological progress. Note that the output/public-capital ratio is dependent on multiple state variables:

$$\sigma\left(K^p, K^g, L, \Phi\right) = \frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N K_i^g} = \frac{\sum_{i=1}^N F_i\left(K_i^p, \Phi K_i^g, L_i\right)}{\sum_{i=1}^N K_i^g} \tag{1.25}$$

where K^p, K^g and L are vectors of private capital, public capital and labor respectively. Moreover, σ is implicitly time dependent since it is constructed from time-varying entities. It follows that the rate of technological progress is

$$\frac{d\Phi}{dt} = \tau\left[\Theta - \sigma\left(K^p, K^g, L, \Phi\right)\right]_+ \tag{1.26}$$

where $\tau \in \Re_+^1$ is an exogenous constant of proportionality and $[.]_+$ is the nonnegative orthant projection operator with the property that $[Q]_+ = \max(0, Q)$ for an arbitrary argument Q.

1.4.4 Criterion Functional and Final Form of the Model

In what follows, we have assumed for simplicity that there is full labor force participation, so that $L_i = P_i$. We wish to maximize the present value of the national income time steam for the time interval $[t_0, t_f]$ expressed as

$$\max J(K^p, K^g, P) = \sum_{i=1}^{N} \int_{t_0}^{t_f} \exp(-\rho t)\, F_i\left(K_i^p, K_i^g, P_i\right) dt \tag{1.27}$$

where $\rho > 0$ is the constant nominal rate of discount and P is presently a vector of regional specific populations:

$$P = (P_i : i \in N)$$

This maximization is to be carried out relative to the dynamics and constraints developed above. Hence, the final form of the model is

$$\max J\left(K^p, K^g, P\right)$$

subject to

$$\frac{dK_i^p}{dt} = (1 - c_i - r)\, F_i\left(K_i^p, \Phi K_i^g, P_i\right) + \eta_i r \sum_{j=1}^{N} F_j\left(K_j^p, \Phi K_j^g, P_j\right) - \delta_p K_i^p \qquad \forall i \in [1, N]$$

and

$$\frac{dK_i^g}{dt} = \nu_i r \sum_{j=1}^{N} F_j\left(K_j^p, \Phi K_j^g, P_j\right) - \delta_g K_i^g \qquad \forall i \in [1, N]$$

$$\frac{dP_i}{dt} = \gamma_i P_i\left(\zeta_i - P_i\right) + \sum_{j \neq i} \lambda_j\left(P_j - P_i\right) \qquad \forall i \in [1, N]$$

$$\frac{d\Phi}{dt} = \tau\left[\Theta - \sigma\left(K^p, K^g, P, \Phi\right)\right]_+$$

$$\sum_{i=1}^{N} (\eta_i + \nu_i) = 1 \qquad \forall i \in [1, N]$$

$$0 \leq \eta_i \leq 1 \qquad \forall i \in [1, N]$$

$$0 \leq \nu_i \leq 1 \qquad \forall i \in [1, N]$$

$$P_i \geq 0 \qquad \forall i \in [1, N]$$

where the shares η_i and v_i for all $i \in [0, N]$, as well as the tax rate r, are the control variables. The state variables are, of course, K_i^p, K_i^g, and P_i for all $i \in [0, N]$, as well as Φ. We have taken the only technological progress to be public capital augmenting, although clearly other options exist. The above model is discussed in more detail in Chapter 7, where a numerical example of it is also presented.

1.5 Dynamic Telecommunications Flow Routing

In telecommunications theory we distinguish between flow routing and flow control. In flow routing we are concerned with the optimal routing of known message demands. In flow control we are concerned with managing demand, including the rejection of message transmission requests. Both classes of problems are amenable to dynamic optimization. For the purpose of providing a preliminary telecommunications example of an optimal control model, we shall presently focus on flow routing.

1.5.1 Assumptions and Notation

The dynamic flow routing problem is concerned with the routing of known and forecast message demands that vary with time in order to minimize the congestion that will be encountered on the telecommunications network. For this problem, we assume that there is a real physical network based on a graph $G(\mathcal{N}, \mathcal{A})$ where $\mathcal{A}$ is a set of directed arcs and $\mathcal{N}$ is a set of nodes. Every arc has associated with it a delay function that derives from a simple model arc latency. That is, the arc delay functions we employ view the traversal time (or latency) experienced by a message packet on arc $a \in \mathcal{A}$, denoted by $D_a[x_a(t)]$, as a function of $x_a(t)$, the message volume on arc a in front of the packet when it enters the arc at time t. The units of this function are delay (time) per unit of flow. We assume that arc delay is always positive, so that

$$D_a[x_a(t)] > 0 \quad \forall a \in \mathcal{A} \tag{1.28}$$

for any argument $x_a \geq 0$. Spillbacks that impact upstream nodes are not considered.

It is convenient to describe a given path (or route) p through the network as a sequence of arcs named in the following fashion:

$$p \equiv \{a_1, a_2, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_{m(p)}\} \tag{1.29}$$

where $m(p)$ is the number of arcs in path p. We will describe the flow dynamics of each arc $a_i \in p$ by

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_i}^p(t) - g_{a_{i-1}}^p(t) \tag{1.30}$$

where

$$\begin{aligned} x_{a_i}^p(t) &= \text{the volume on arc } a_i \text{ due to flow on path } p \text{ at time } t \\ g_{a_i}^p(t) &= \text{the flow exiting arc } a_i \text{ of path } p \text{ at time } t \\ g_{a_{i-1}}^p(t) &= \text{ the flow entering arc } a_{i-1} \text{ of path } p \text{ at time } t \end{aligned}$$

Furthermore, we recognize that $g_{a_0}^p$ is the flow exiting the origin node of path p. We give $g_{a_0}^p$ the special symbol h_p and refer to it as the departure rate into path p, or more simply as the *flow on path* p when this latter name is not confusing.

Of course, volume on a given arc is the sum of contributions from the paths traversing that arc; it is given for every $a \in \mathcal{A}$ by

$$x_a = \sum_{p \in \mathcal{P}} \delta_{ap} x_a^p \tag{1.31}$$

where

$$\delta_{ap} = \begin{cases} 1 & \text{if } a \in p \\ 0 & \text{if } a \notin p \end{cases}$$

and $\mathcal{P}$ denotes the set of paths connecting origin-destination (OD) pairs of the network. We also use $\mathcal{N}_O$ to denote the set of nodes from which message traffic originates and $\mathcal{N}_D$ to denote the set of nodes for which traffic is destined. We also use $\mathcal{W} = \{(i, j) : i \in \mathcal{N}_O, j \in \mathcal{N}_D\}$ to denote the set of origin-destination pairs between which message traffic moves. Furthermore, $\mathcal{P}_{ij}$ will be the set of paths connecting OD pair $(i, j) \in \mathcal{W}$, so that

$$\mathcal{P} = \bigcup_{(i,j) \in \mathcal{W}} \mathcal{P}_{ij}$$

describes the set of all network paths.

1.5.2 Flow Propagation Mechanism

We now develop a mechanism to describe the physical propagation of flows through the network. To this end we introduce the concept of the *arc exit time function*. To understand the exit time function, let t_e be the time at which flow exits the ith arc of path p when departure from the origin of that path has occurred at time t_d. The relationship of these two instants of time is expressed as

$$t_e = \tau_{a_i}^p(t_d) \tag{1.32}$$

and we call $\tau_{a_i}^p(.)$ the exit time function for arc a_i of path p. The inverse of the exit time function is written as

$$t_d = \theta_{a_i}^p(t_e) \tag{1.33}$$

and describes the time of departure t_d from the origin of path p for flow which exits arc a_i of that path at time t_e. Consequently, the following identity must hold

$$t = \theta_{a_i}^p\left(\tau_{a_i}^p(t)\right) \tag{1.34}$$

for all time t for which flow behavior is being modeled. It then follows immediately that the total traversal time for path p can be articulated in terms of the final exit time function and the departure time:

$$D_p(t) = \sum_{i=1}^{m(p)} \left[\tau_{a_i}^p(t) - \tau_{a_i-1}^p(t)\right] = \tau_{a_{m(p)}}^p(t) - t \tag{1.35}$$

when departure from the origin of path p is at time t.

A further consequence of the assumed model of arc delay is

$$\tau_{a_1}^p = t + D_{a_1}\left[x_{a_1}(t)\right] \qquad \forall\ p \in \mathcal{P} \tag{1.36}$$

$$\tau_{a_i}^p = \tau_{a_{i-1}}^p(t) + D_{a_i}\left[x_{a_i}\left(\tau_{a_{i-1}}^p(t)\right)\right] \qquad \forall\ p \in \mathcal{P},\ i \in [2, m(p)] \tag{1.37}$$

Differentiating (1.36) and (1.37) with respect to time gives

$$\frac{d\tau_{a_1}^p(t)}{dt} = 1 + D'_{a_1}\left[x_{a_1}(t)\right]\frac{dx_{a_1}(t)}{dt} \qquad \forall\ p \in \mathcal{P} \tag{1.38}$$

$$\frac{d\tau_{a_i}^p(t)}{dt} = \left[1 + D'_{a_i}\left[x_{a_i}\left(\tau_{a_{i-1}}^p(t)\right)\right]\frac{dx_{a_1}\left[\tau_{a_{i-1}}^p(t)\right]}{d\tau_{a_{i-1}}^p(t)}\right]\frac{d\tau_{a_{i-1}}^p(t)}{dt} \tag{1.39}$$

$$\forall\ p \in \mathcal{P},\ i \in [2, m(p)]$$

where we have again used the chain rule and the prime superscript denotes total differentiation with respect to the associated function argument. Evidently, expressions (1.30), (1.36), and (1.38) are easily combined to yield

$$g_{a_1}\left(t + D_{a_1}\left[x_{a_1}(t)\right]\right)\left(1 + D'_{a_1}\left[x_{a_1}(t)\right]\dot{x}_{a_1}(t)\right) = h_p(t) \tag{1.40}$$

where the overdot refers to a total time derivative. Proceeding inductively from this last result with the guidance of (1.39), we obtain

$$g_{a_i}^p\left(t + D_{a_i}\left[x_{a_i}(t)\right]\right)\left(1 + D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}(t)\right) = g_{a_{i-1}}^p(t) \tag{1.41}$$

$$\forall\ p \in \mathcal{P},\ i \in [2, m(p)]$$

Expressions (1.40) and (1.41) are *proper flow progression constraints* derived in a fashion that makes them completely consistent with the assumed model of arc delay. Note that these constraints involve a state-dependent time lag $D_{a_i}\left[x_{a_i}(t)\right]$ but make *no explicit reference to the exit time functions and their inverses*. We will subsequently use (1.40) and (1.41) as constraints for dynamic flow routing in order to assure physically meaningful flow.

1.5.3 Path Delay Operators

The preceding development allows us to determine closed-form path delay operators that tell us the delay experienced by a message packet transmitted at time t and encountering traffic conditions x. In particular, we note that the recursive relationships (1.36) and (1.37) lead to the operators

$$D_p(t,x) \equiv \sum_{i=1}^{m(p)} \Phi_{a_i}(t,x) \tag{1.42}$$

for traffic conditions

$$x \equiv \left(x_{a_i}^p : p \in [1, |\mathcal{P}|], i \in [1, m(p)]\right)$$

where the $\Phi_{a_i}(t,x)$ are arc delay operators obeying

$$\left.\begin{aligned}
&\Phi_{a_1}(t,x) = D_{a_1}\left[x_{a_1}(t)\right] > 0\\
&\Phi_{a_2}(t,x) = D_{a_2}\left[x_{a_2}\left(t+\Phi_{a_1}\right)\right] > 0\\
&\Phi_{a_3}(t,x) = D_{a_3}\left[x_{a_3}\left(t+\Phi_{a_1}+\Phi_{a_2}\right)\right] > 0\\
&\vdots\\
&\Phi_{a_i}(t,x) = D_{a_i}\left[x_{a_i}\left(t+\sum_{j=1}^{i-1}\Phi_{a_j}\right)\right] > 0
\end{aligned}\right\} \tag{1.43}$$

We also introduce the arrival penalty operator

$$\Pi\left[t + D_p(t,x) - T_A\right] \tag{1.44}$$

where T_A is the prescribed fixed arrival time. The arrival penalty operator has the properties

$$t + D_p(t,x) > T_A \Longrightarrow \Pi\left[t + D_p(t,x) - T_A\right] > 0 \tag{1.45}$$

$$t + D_p(t,x) < T_A \Longrightarrow \Pi\left[t + D_p(t,x) - T_A\right] > 0 \tag{1.46}$$

$$t + D_p(t,x) = T_A \Longrightarrow \Pi\left[t + D_p(t,x) - T_A\right] = 0 \tag{1.47}$$

for every path $p \in \mathcal{P}$. Consequently, the effective delay operator for each path $p \in \mathcal{P}$ is

$$\Psi_p(t,x) = D_p(t,x) + \Pi\left[t + D_p(t,x) - T_A\right] > 0 \tag{1.48}$$

Note carefully that the path delay operators have been expressed as a closed form in x requiring as *a priori* knowledge only the arc delay functions. This is a powerful result: expression (1.48) tells us literally how to look into the future to model the delay of message packets we are transmitting in the present. In that path flows determine arc exit rates and arc volumes, we may also use the notation

$$\Psi_p(t,h) \quad \forall p \in \mathcal{P}$$

to denote the effective delay operators.

1.5.4 Dynamic System Optimal Flows

We imagine that there is a central authority that manages the network, setting message transmission rates and determining message routes. Therefore, we are interested in minimizing total system wide delay for the full network based on the graph $G(\mathcal{N},\mathcal{A})$ over the period $\left[t_0, t_f\right]$, expressed as

$$\min J_1 = \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p[t, h(t)]\, h_p(t)\, dt \tag{1.49}$$

where $t = t_0$ is the earliest allowed transmission time and t_f is the transmission time horizon. That is, no message traffic can be transmitted prior to time t_0 and all messages must have been transmitted by time t_f. It is easy to introduce notation to allow for different classes of messages with distinct arrival times, but we refrain from doing so to keep the notation simple.

1.5.5 Additional Constraints

There will be two types of demand: scheduled demand and instantaneous demand. Scheduled demand is known *a priori*, by which we mean at or prior to time t_0. Scheduled demand is serviced at the convenience of the controller so long as it is serviced prior to the end of the planning horizon. By contrast, instantaneous demand must be serviced at the time it arises. We denote the fixed, scheduled demand for origin-destination pair $(i,j) \in \mathcal{W}$ by Q_{ij}; the instantaneous demand at time $t \in \left[t_0, t_f\right]$ for the same OD pair will be $R_{ij}(t)$. Both Q_{ij} and $R_{ij}(t)$ are exogenously determined and known exactly. This is somewhat unrealistic as demand forecasts will be subject to error so that each $R_{ij}(t)$ must necessarily have a stochastic error. Given the deterministic focus of this book, we ignore this bit of reality so that the

basic structure of the model can be outlined with the minimum amount of notation. Therefore, we impose the following flow generation and conservation constraints:

$$\sum_{p \in \mathcal{P}_{ij}} h_p(t) \geq R_{ij}(t) \tag{1.50}$$

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} + \int_{t_0}^{t_f} R_{ij}(t)\, dt \tag{1.51}$$

for every defined $(i, j) \in \mathcal{W}$. Note that constraint (1.50) requires that instantaneous demand be serviced at the time it arises; there may be a technological lag which requires that (1.50) be replaced by the constraint

$$\sum_{p \in \mathcal{P}_{ij}} h_p(t + \zeta) \geq R_{ij}(t) \tag{1.52}$$

or all $(i, j) \in \mathcal{W}$ where ζ is a fixed constant time lag reflecting the time needed to respond to demand. A still richer model can be constructed by assuming that the lag ζ is origin-destination specific and depends on the prevailing congestion level.

We also impose the nonnegativity restrictions

$$x \geq 0 \qquad g \geq 0 \qquad h \geq 0 \tag{1.53}$$

where

$$x \equiv \left(x_{a_i}^p : p \in [1, |\mathcal{P}|], i \in [1, m(p)]\right) \tag{1.54}$$

$$g \equiv \left(g_{a_i}^p : p \in [1, |\mathcal{P}|], i \in [1, m(p)]\right) \tag{1.55}$$

$$h \equiv \left(h_p : p \in [1, |\mathcal{P}|]\right) \tag{1.56}$$

are the vectors of state and control variables. Individual arcs may be assumed to have hard capacity constraints on either volumes, entry flows or exit flows as is appropriate for the network being studied. Arc volume capacity constraints can, of course, also be reflected in the arc delay functions. For example, the function

$$D_a = A_a + \frac{B_a}{K_a - x_a}$$

ensures that arc volume x_a can never exceed the known, fixed capacity K_a. In light of the preceding development, if we employ arc delay functions with embedded capacities, the set

$$\Lambda_1 = \{(x, h, g) : (1.40), (1.41), (1.50), (1.51), \text{ and } (1.53) \text{ hold}\} \tag{1.57}$$

describes the feasible region of the omniscient controller for flow routing.

1.5.6 Final Form of the Model

As a consequence of the preceding development, the telecommunications flow routing problem in a deterministic setting may be stated as

$$\left.\begin{array}{ll} \min J_1 = \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p\left[t, x\left(t\right)\right] h_p\left(t\right) dt & \\ s.t. \qquad\qquad (x, h, g) \in \Lambda_1 & \end{array}\right\} \tag{1.58}$$

This model is a direct extension of the quasistatic flow routing problem defined by Bertsekas and Gallager (1992) to a dynamic setting. Traditional necessary conditions for optimal control problems cannot be employed because (1.58) has embedded time shifts which are state-dependent. Friesz et al. (2001) have developed necessary conditions for this class of problems. These conditions, when applied to (1.58), reveal optimal control policies that equate marginal costs on open paths and are a synthesis of bang-bang and singular controls. Friesz et al. (2004) have used projection methods to solve this problem. Friesz and Mookherjee (2006) have proposed a fixed-point algorithm for models of this type.

1.6 Brief History of Dynamic Games

The latter half of the twentieth century saw impressive achievements in the modeling, analysis, and computation of competitive static equilibria of noncooperative games, as was underscored by the joint award of a Nobel Prize in economics to John Nash, John Harsanyi, and Reinhard Selten in 1993 for their fundamental work on mathematical games and the relationship of games to equilibrium and optimization. Mathematicians, game theorists, operations researchers, economists, biologists, and engineers have all employed noncooperative mathematical games and the notion of equilibrium to model virtually every kind of competition. In particular, noncooperative game-theoretic models have been successfully employed to study economic competition in the marketplace, highway and transit traffic in the presence of congestion, wars, and both intra- and interspecies biological competition. One of the key developments that has made such diverse applications practical is our ability to compute static game-theoretic equilibria as solutions of appropriately defined variational inequalities, nonlinear complementarity problems, and fixed-point problems.

In many applications, intermediate disequilibrium states of mathematical games are intrinsically important. When this is the case, disequilibrium adjustment processes[4] must be articulated, thereby, forcing explicit consideration of time.

[4] Such adjustment processes are typically expressed as difference equations or differential equations. Sometimes it is necessary to use a mixture of both types of dynamics; the adjustment processes are then referred to as *differential-difference equations*.

Also, sometimes flow conservation constraints must be formulated in such a way that they constitute equilibrium dynamics; this usually occurs when a crucial state variable must be differentiated in order to form a flow conservation law. As a consequence, the modeling of competitive equilibria and disequilibria frequently involves dynamic or differential games. While great progress has been made in modeling and computation of static equilibria or steady states of competitive systems, game-theoretic disequilibria and moving equilibria[5] are relatively uninvestigated by comparison.

The main body of technical literature relevant to game-theoretic disequilibria and moving equilibria is that pertaining to so-called *differential games*, a field of inquiry widely held to have been originated by Isaacs (1965). Although a rather substantial body of literature known as *dynamic game theory* has evolved from the work of Isaacs (1965), that literature is characterized by an emphasis on the relationship of such games to dynamic programming and to the Hamilton-Jacobi-Bellman partial differential equation.[6] A consequence of this classical point of view is that full use of the mathematical apparatus of variational inequalities, discovered originally in the context of certain free boundary-value problems in mathematical physics, has not occurred in the study of dynamic games. By contrast, in the last fifteen years, variational inequalities have become the formalism of choice for applied game theorists and computational economists solving various static equilibrium models of competition. The "hole" in the dynamic game theory literature owing to this failure to fully exploit the variational inequality perspective is significant, for variational inequalities substantially simplify the study of existence and uniqueness. A variational inequality perspective for infinite-dimensional dynamic games also leads directly to function space equivalents of the standard finite-dimensional algorithmic philosophies of feasible direction and projection familiar from nonlinear programming.

As examples of non-cooperative dynamic games expressible as differential variational inequalities, we consider three problems:

1. dynamic traffic equilibrium;
2. dynamic oligopolistic network competition; and
3. competitive dynamic revenue management and pricing.

1.7 Dynamic User Equilibrium for Vehicular Networks

Because of the increased importance of information technology for road traffic, much attention has been devoted in the last few years to the problem of predicting time-varying (dis)equilibrium flows on passenger car networks. A predictive ability

[5] We define *moving equilibrium* in a subsequent chapter; however, it suffices at this juncture to think of a moving equilibrium as a trajectory of decison variables that maintains the same "balance" among those variables throughout time, although the variables themselves are time-varying.

[6] See, for example, Basar and Olsder (1998).

of this type is necessary to provide data unavailable from real-time sensors and to forecast future traffic conditions in order to construct route advisories. A dynamic network user equilibrium (DUE) model of road traffic is quite similar in some respects to that of the dynamic system optimal (DSO) model presented above for telecommunications flow routing. The notation, for one thing, is identical. The most fundamental difference lies in the fact that in DUE the DSO objective function is replaced with an appropriate variational inequality to capture the noncooperative gaming behavior of drivers. In fact, we are led to posit the following formulation of the dynamic network user equilibrium problem: find $(x^*, g^*, h^*) \in \Gamma$ such that

$$\left\langle \Psi\left(t, x^*\right), h - h^* \right\rangle \equiv \sum_{p \in P} \int_{t_0}^{t_f} \Psi_p\left(t, x^*\right) \left[h_p\left(t\right) - h_p^*\left(t\right)\right] dt \geq 0 \tag{1.59}$$

for all $(x, g, h) \in \Gamma$, where Γ is the set of all (x, g, h) obeying

$$\frac{dx_{a_1}^p\left(t\right)}{dt} = h_p\left(t\right) - g_{a_1}^p\left(t\right) \qquad \forall p \in P \tag{1.60}$$

$$\frac{dx_{a_i}^p\left(t\right)}{dt} = g_{a_{i-1}}^p\left(t\right) - g_{a_i}^p\left(t\right) \qquad \forall\ p \in P, i \in [2, m\left(p\right)] \tag{1.61}$$

$$(x, g, h) \in \Omega \tag{1.62}$$

$$x\left(0\right) = x^0 \tag{1.63}$$

Bernstein et al. (1993) were the first to propose a DUE model structure like (1.59) through (1.63). We now refer to such problems as *differential variational inequalities*, a type of problem we define formally in Chapter 6. The main motivation for offering formulation (1.59) through (1.63) is that the variational inequality (1.59) is known from Friesz et al. (1993) to describe a Nash-like dynamic equilibrium, while (1.60) through (1.63), as mentioned in our prior discussion of telecommunications flow routing, are known to be valid constrained dynamics. However, for the reader unfamiliar with such notions, formulation (1.59) through (1.63) is conjectural; a formal demonstration that its solutions will in fact be dynamic network user equilibria is required. We postpone that analysis until we have derived and mastered the necessary conditions for optimal control problems and differential variational inequalities.

1.8 Dynamic Oligopolistic Network Competition

We consider in this section a version of the dynamic oligopolistic network competition problem due to Friesz et al. (2006). The oligopolistic firms of interest, embedded in a network economy, are in oligopolistic game-theoretic competition described by a Nash equilibrium. That equilibrium includes dynamics that describe

the trajectories of inventories/backorders and correspond to flow conservation for each firm at each node of the network of interest. The oligopolistic firms, acting as shippers, compete as price takers in the market for physical distribution services, which is perfectly competitive due to its involvement in other markets of the network economy. The time scale we consider is neither short nor long, but rather of sufficient length to allow output and shipping pattern adjustments, yet not long enough for firms to relocate or enter or leave the network economy.

1.8.1 Notation

We employ the notation used in Miller et al. (1996), augmented to handle temporal considerations. In particular, time is denoted by the continuous scalar $t \in \Re^1_+$ and the analysis period by $\left[t_0, t_f\right] \subseteq \Re^1_+$ where $t_0 < t_f$. There are several sets important to articulating a model of oligopolistic competition on a network; these are as follow: $\mathcal{F}$ for firms, $\mathcal{A}$ for directed arcs, $\mathcal{N}$ for nodes, and $\mathcal{W}$ for origin-destination (OD) pairs. Subsets of these sets are formed as is meaningful by using the subscript f for a specific firm, i for a specific node, and w for a specific OD pair.

Each firm under consideration controls production rates q^f, allocations of output to meet demand c^f, and shipping rates s^f. Inventories I^f are state variables determined by the controls. In particular, c^f, q^f, s^f, and I^f may be concatenated to form the following:

$$
\begin{aligned}
c &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \\
q &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \\
s &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{W}|\times|\mathcal{F}|}
\end{aligned}
$$

$$
I\,(c,q,s) : \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \times \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \times \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{W}|\times|\mathcal{F}|} \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|}
$$

where $L^2\left[t_0, t_f\right]$ is the space of square-integrable functions and $\mathcal{H}^1\left[t_0, t_f\right]$ is a Sobolev space for the real interval $\left[t_0, t_f\right] \in \Re^1_+$.

1.8.2 Extremal Problems and the Nash Game

Each firm has an objective of maximizing net profit expressed as revenue less cost and taking the form of an operator acting on allocations of output to meet demands, production rates, and shipment patterns. Let $\pi_i(.,t)$, for each node (market) $i \in \mathcal{N}$,

be the inverse demand function for the homogeneous good that is produced by all firms. For each $f \in \mathcal{F}$, net profit is

$$\Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f}) = \int_{t_0}^{t_f} e^{-\rho t} \left\{ \sum_{i \in \mathcal{N}} \pi_i \left(\sum\nolimits_{g \in \mathcal{F}} c_i^g, t \right) c_i^f - \sum_{i \in \mathcal{N}_f} V_i^f(q^f, t) - \sum_{w \in \mathcal{W}_f} r_w(t)\, s_w^f - \sum_{i \in \mathcal{N}} \psi_i^f \left(I_i^f, t \right) \right\} dt \tag{1.64}$$

where $\rho \in \Re^1_{++}$ is a constant nominal rate of discount, $r_w \in \Re^1_{++}$ is the freight rate (tariff) charged per unit of flow s_w for origin-destination (OD) pair $w \in \mathcal{W}_f$, ψ_i^f is firm f's inventory cost at node i, and I_i^f is the inventory/backorder of firm f at node i. In (1.64), c_i^f is the allocation of the output/inventory of firm $f \in \mathcal{F}$ at node $i \in \mathcal{N}$ to consumption at that node. Our formulation is in terms of flows, so we employ the inverse demand functions $\pi_i(c_i, t)$ where

$$c_i = \sum_{g \in \mathcal{F}} c_i^g$$

is the total allocation of output and/or inventory to consumption for node $i \in \mathcal{N}$. Furthermore, q_i^f is the output of firm $f \in \mathcal{F}$ at node $i \in \mathcal{N}$. Also $V_i^f(q, t)$ is the variable cost of production for firm $f \in \mathcal{F}$ at node $i \in \mathcal{N}$. Note that $\theta_f(c^f, q^f, s^f; c^{-f}, q^{-f})$ is a functional that is completely determined by the controls c^f, q^f and s^f when non-own allocations to consumption and non-own production rates

$$c^{-f} \equiv (c^{f'} : f' \neq f)$$
$$q^{-f} \equiv (q^{f'} : f' \neq f)$$

are taken as exogenous data by firm f. In expression (1.64), the first term of the functional $\theta_f(c^f, q^f, s^f; c^{-f}, q^{-f})$ is the firm's revenue; the second term is the firm's cost of production; the third term is the firm's shipping costs; and the last term is the firms's inventory or holding cost.

We also impose the terminal time inventory constraints

$$I_i^f(t_f) = \tilde{K}_i^f \quad \forall\, f \in \mathcal{F}, i \in \mathcal{N}_f \tag{1.65}$$

where $\tilde{K}_i^f \in \Re^1_+$ is exogenous. In the event a specific $\tilde{K}_i^f$ is strictly positive, we need to include in the objective functional an associated penalty for or salvage

value of residual inventory. All consumption, production and shipping variables are nonnegative and bounded from above; that is

$$C^f \geq c^f \geq 0 \tag{1.66}$$
$$Q^f \geq q^f \geq 0 \tag{1.67}$$
$$S^f \geq s^f \geq 0 \tag{1.68}$$

where

$$C^f \in \Re_{++}^{|\mathcal{F}|}$$
$$Q^f \in \Re_{++}^{|\mathcal{F}|}$$
$$S^f \in \Re_{++}^{|\mathcal{W}_f|}$$

are vectors of exogenous parameters. Constraints (1.66), (1.67), and (1.68) are recognized as pure control constraints, while (1.65) are terminal conditions for the state variables. Naturally

$$\Omega_f = \left\{\left(c^f, q^f, s^f\right) : (1.66), (1.67), (1.68)\right\}$$

is the set of feasible controls.

Firm f solves an optimal control problem to determine its production q^f, allocation of production to meet demand c^f, and shipping pattern s^f – thereby also determining inventory I^f via dynamics we articulate momentarily – by maximizing its profit functional $\Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f})$ subject to inventory dynamics expressed as flow balance equations and pertinent production and inventory constraints. The inventory dynamics for firm $f \in \mathcal{F}$, expressing simple flow conservation, obey

$$\frac{dI_i^f}{dt} = E_i^f \quad \forall\, i \in \mathcal{N}_f \tag{1.69}$$

$$I_i^f(t_0) = K_i^f \quad \forall\, i \in \mathcal{N}_f \tag{1.70}$$

where E_i^f is excess goods flow obeying

$$E_i^f \equiv q_i^f + \sum_{w \in \mathcal{W}_i^d} s_w^f - \sum_{w \in \mathcal{W}_i^o} s_w^f - c_i^f \quad \forall\, f \in \mathcal{F}, i \in \mathcal{N}_f$$

while the $K_i^f \in \Re_{++}^1$ are exogenous parameters, $\mathcal{W}_i^d$ is the set of OD pairs with destination node i, and $\mathcal{W}_i^o$ is the set of OD pairs with origin node i. Consequently

$$I(c, q, s) = \arg\left\{\frac{dI_i^f}{dt} = E_i^f, \quad I_i^f(t_0) = K_i^f, \quad I_i^f\left(t_f\right) = \tilde{K}_i^f \right.$$
$$\left. \forall\, f \in \mathcal{F}, i \in \mathcal{N}_f\right\}$$

where we implicitly assume that the dynamics have an unique solution for all feasible controls.

With the preceding development, we note that firm f's problem is: with the c^{-f} and q^{-f} as exogenous inputs, compute c^f, q^f and s^f (thereby finding I^f) in order to solve the following extremal problem:

$$\left.\begin{array}{ll} \max & \Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f}) \\ \text{subject to} & \left(c^f, q^f, s^f\right) \in \Omega_f \end{array}\right\} \forall f \in \mathcal{F} \tag{1.71}$$

where

$$\Omega_f = \left\{\left(c^f, q^f, s^f\right) : (1.65), (1.66), (1.67), (1.68) \text{ hold}\right\}$$

also for all $f \in \mathcal{F}$. That is, each firm is a Nash agent that knows and employs the current instantaneous values of the decision variables of other firms to make its own non-cooperative decisions.

1.8.3 Differential Variational Inequality Formulation

Note that (1.71) defines a non-cooperative game expressed as a set of coupled optimal control problems, one for each firm $f \in \mathcal{F}$. Its solution is a so-called Nash equilibrium, a notion we study in some detail in subsequent chapters. As demonstrated formally by Friesz et al. (2006), solutions of the following variational inequality, when they exist, are Nash equilibria for the above non-cooperative, open-loop game: find $\left(c^{f*}, q^{f*}, s^{f*}\right) \in \Omega$ such that

$$\sum_{f\in\mathcal{F}} \int_{t_0}^{t_f} \left[\sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial c_i^f}\left(c_i^f - c_i^{f*}\right) + \sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial q_i^f}\left(q_i^f - q_i^{f*}\right) + \sum_{w\in\mathcal{W}_f} \frac{\partial \Phi_f^*}{\partial s_w^f}\left(s_w^f - s_w^{f*}\right) dt \right] \geq 0 \tag{1.72}$$

for all $(c, q, s) \in \Omega$, where

$$\begin{aligned} \Phi_f\left(c^f, q^f, s^f, I^f; c^{-f}, q^{-f}; t\right) = e^{-\rho t} \Bigg\{ & \sum_{i\in\mathcal{N}} \pi_i\left(\sum\nolimits_{g\in\mathcal{F}} c_i^g, t\right) c_i^f \\ & - \sum_{i\in\mathcal{N}_f} V_i^f(q,t) - \sum_{w\in\mathcal{W}_f} r_w(t)\, s_w^f \\ & - \sum_{i\in\mathcal{N}} \psi_i^f\left(I_i^f(c,q,s), t\right) \Bigg\} \end{aligned}$$

$$\Phi_f^* = \Phi_f \left(c^{f*}, q^{f*}, s^{f*}, I^{f*}; c^{-f*}, q^{-f*}; t \right)$$
$$\Omega = \prod_{f \in \mathcal{F}} \Omega_f$$

The variational inequality (1.72) is a convenient way to express dynamic oligopolistic network competition.

1.9 Revenue Management and Nonlinear Pricing

Revenue management (RM), also known as *revenue optimization*, is a relatively new field of inquiry. Revenue management is concerned with extracting all unused willingness to pay from consumers of services provided by individual firms who compete with one another for market shares. Growth of the field of revenue management was boosted by the deregulation of domestic and international airlines in the late 1970's. Airlines, as well as other service firms, exercise quantity-based RM techniques by controlling the resources sold during a booking period at a prespecified price. In contrast, retailers are said to use price-based RM when they exploit price as an instrument to control demand over the selling period. See McGill and van Ryzin (1999) for a detailed survey of research on mathematical models and quantitative tools for RM from 1970 through the late 1990s.

1.9.1 The Decision Environment

In this section we consider service firms that provide differentiated, nonsubstitutable services, set prices for their services, may decline some of the booking requests at any given time, face cancellations (with full or partial refunds) and have finite supplies of resources. The demand is assumed to be known with certainty, a severe assumption mandated by the introductory and deterministic focus of this book. Stochastic extensions of the model reported here are found in Mookherjee and Friesz (2008). Each service firm has to decide *both* how to allocate its resources (quantity-based RM) and how to set prices for its services (price-based RM) as it seeks to maximize its revenue. Such a decision environment differs from that faced by discount airlines mainly in the absence of demand uncertainty.

Consider an oligopoly of abstract service providers. Each firm provides a set of services (products). Each network product may be viewed as a bundle of resources sold with certain terms of purchase and restrictions at a given price. As already noted, these services are nonsubstitutable and differentiated. Furthermore, all firms have finite resources. The booking period is is expressed in N discrete time periods $t \in [0, N-1]$. At the beginning of a discrete period, firms set service prices and quantities for sale in that period. The set of all firms is denoted by $\mathcal{F}$, while $\mathcal{S}$ denotes the set of services provided by each form. Furthermore, $\mathcal{C}$ will denote the

set of resources firms use to provide services; $\mathcal{C}_i$ is the set of resources used to provide service $i \in \mathcal{S}$; and $\mathcal{S}_j$ is the set of services that utilize resource $j \in \mathcal{C}$. Also, A is the resource-service incidence matrix, and $|\mathcal{C}|$ is the cardinality of $\mathcal{C}$, with analogous definitions for other sets.

There is also notation for a number of parameters that must be introduced. The minimum price that firm f can charge for service $i \in \mathcal{S}$ will be denoted by $p^f_{i,\min}$, and $p^f_{i,\max}$ will be the maximum price that firm f can charge for service $i \in \mathcal{S}$. Also, there will be a strictly positive minimum allowed level of service denoted by $u_{\min} \in \Re^1_{++}$. The capacity of firm $f \in \mathcal{F}$ for resource type $j \in \mathcal{C}$ will be named K^f_j, while ρ^f_t will denote the cancellation rate for firm $f \in \mathcal{C}$ at the end of period t.

We turn now to the variables and functions we will employ for our illustrative revenue management model. In particular, $p^f_{i,t}$ will be the price for service $i \in \mathcal{S}$ charged by firm $f \in \mathcal{F}$ in time period $t \in [0, N-1]$, while $x^f_{j,t}$ will be the allocation of resource type $j \in \mathcal{C}$ by firm $f \in \mathcal{F}$, also in in time period $t \in [0, N-1]$. The demand realized by firm $f \in \mathcal{F}$ for service $i \in \mathcal{S}$ will be denoted by $D^f_{i,t}(p_{i,t})$ in time period $t \in [0, N-1]$. The refund by firm $f \in \mathcal{F}$ for cancelling resource $j \in \mathcal{C}$, also in time period $t \in [0, N-1]$, will be $R^f_{j,t}(.)$ while $\Psi^f_N(.,.)$ will be denial-of-service cost for firm $f \in \mathcal{F}$ at the end of period N. The vector of decision variables for service i provided by firm f are

$$p^f_t = \left(p^f_{i,t} : i \in \mathcal{S}\right)$$

which concatenates to

$$p^f = \left(p^f_t : t \in [0, N-1]\right)$$

The pricing decision variables of firm f's competitors for period t are denoted by the vector

$$p^{-f}_t = \left(p^g_t : g \in \mathcal{F} \backslash f\right),$$

The state variables for each firm f are the vectors of cumulative allocations of resources

$$x^f_t = \left(x^f_{j,t} : j \in \mathcal{C}\right)$$

for period t. The network we are interested in has $|\mathcal{C}|$ resources and the firm provides $|\mathcal{S}|$ different services. Each network product is a combination of a bundle of the $|\mathcal{C}|$ resources sold with certain terms of purchase and restrictions at a given price. The resource-service *incidence matrix*, $A = \left[a_{ij}\right]$ is a $|\mathcal{C}| \times |\mathcal{S}|$ matrix where

$$a_{ij} = \begin{cases} 1 & \text{if resource } i \text{ is used for service } j \\ 0 & \text{otherwise} \end{cases}$$

Thus, the jth column of A, denoted by A_j, is the *incidence vector* for service j; while the ith row, denoted by A^i, has unity in column j provided service j utilizes resource i.

1.9.2 The Role of Denial-of-Service Costs and Refunds

Typically, service providers are allowed to cancel scheduled services, which are allocated resources. In particular, cancellations are assumed to occur at the rate ρ_t^f for each firm $f \in \mathcal{F}$ and discrete-time periods $t \in [0, N-1]$. Such cancellations require refunds expressed as

$$R_t^f \left(\rho_t^f \cdot x_t^f \right) \tag{1.73}$$

for firm $f \in \mathcal{F}$ in period $t \in [0, N-1]$. Refunds $R_t^f(\cdot)$ should monotonically increase with x_t^f and ρ_t^f, and decrease with time t; such qualitative behavior reflects the potential for cancellation fees to increase as the end of the booking period is approached. Denial of service must necessarily involve loss of goodwill on the part of customers toward service providers. These denial-of-service costs are calculated at the end of the booking period and involve the comparison of resources delivered to actual capacity. Denial-of-service costs are expressed as

$$\Psi_N^f \left(x_N^f, K^f \right) \tag{1.74}$$

for firm $f \in \mathcal{F}$, where of course

$$K^f = \left(K_j^f : j \in \mathcal{C} \right) \tag{1.75}$$

is the vector of actual capacities. If demand, which is a function of final allocation at the terminal time, is less than or equal to the physical capacity, then no denial-of-service cost is incurred; otherwise denial-of-service costs increase monotonically with excess demand.

1.9.3 Firms' Extremal Problem

With the rival firms' prices p_t^{-f} taken as exogenous to firm $f \in \mathcal{F}$'s discrete-time optimal control problem and yet endogenous to the overall model, firm $f \in \mathcal{F}$ computes its prices p_t^f and allocation of resources u_t^f in order to maximize net revenue generated throughout the booking period subject to pertinent constraints:

$$\max_{p^f, u^f} J \left(p^f ; p^{-f} \right) = -\Psi_N^f \left(x_N^f, K^f \right) - \sum_{t=0}^{N-1} R_t^f \left(\rho_t^f \cdot x_t^f \right) + \sum_{t=0}^{N-1} p_t^f \cdot D_t^f \left(p_t \right) \tag{1.76}$$

subject to

$$x_{t+1}^f = x_t^f + A \cdot D_t^f(p_t) - \rho_t^f \cdot x_t^f \quad t = 0, \ldots, N-1 \tag{1.77}$$

$$x_0^f = 0 \tag{1.78}$$

$$p_{\min}^f \le p_t^f \le p_{\max}^f \quad t = 0, \ldots, N-1 \tag{1.79}$$

$$D_t^f(p_t) \ge u_{\min} \quad t = 0, \ldots, N-1 \tag{1.80}$$

The first two terms on the righthand side of (1.76) are the denial-of-service costs and total refunds, respectively. These are subtracted from total revenue generated to give net revenue generated in the booking period. As in a typical RM industry, there is no salvage value of unsold resources at the end of the horizon. Constraints (1.77) are definitional dynamics that describe the net rate of resource commitment. Of course (1.78) is an initial condition that states no resources are committed at the start of the booking period. Service prices are bounded from above and below as in (1.79). Constraints (1.80) serve to bound realized demand away from zero; without this constraint, it is possible for a service provider to offer no service in one or more periods yet to set prices, which is implausible. Like the dynamic user equilibrium problem, the family of extremal problems considered above may be recast as a differential variational inequality, as we shall see in Chapter 10 where a detailed numerical example of a similar model is presented.

1.10 The Material Ahead

In the chapters that follow, the mathematical foundations of dynamic optimization and of differential non-cooperative games are presented. The presentation involves a level of mathematical abstraction appropriate for graduate students and researchers in economics, operations research, industrial engineering, other branches of engineering, and applied mathematics. Although the presentation is unremittingly mathematical, in principle only a background in elementary calculus and linear algebra is required. Prior exposure to mathematical programming, optimal control theory, functional analysis, and measure theory is not mandatory. However, the reader with such prior exposure will be able to move through the material presented much more quickly.

A distinguishing feature of this book is its emphasis on computation. All readers, regardless of prior mathematical preparation, should take the time to work through the computational examples. Doing so will lead to a much deeper understanding of the relationships of infinite-dimensional mathematical programming, optimal control theory, variational inequalities, and differential, non-zero sum games. In particular, the reader who does study the computational examples provided herein will be able to proceed much more rapidly in the development of custom software to solve the dynamic optimization problems and dynamic games that are the focus of his/her own research.

In the chapters ahead, after optimality conditions and algorithms are studied in the abstract, several dynamic modelling applications are presented. It should be noted that the dynamic applications studied in later chapters overlap, by intention, some of the illustrative applications presented in earlier sections of Chapter 1. The applications presented include numerical illustrations of various algorithms and run the gamut from inventory theory to supply chains to traffic assignment to revenue management. Hopefully, there is at least one application of interest to every reader.

List of References Cited and Additional Reading

Arrow, K. J. and M. Kurz (1970). *Public Investment, the Rate of Return, and Optimal Fiscal Policy*. Baltimore: The Johns Hopkins University Press.

Bagchi, A. (1984). *Stackelberg Differential Games in Economic Models*. Berlin: Springer-Verlag.

Basar, T. and G. J. Olsder (1998). *Dynamic Noncooperative Game Theory*. Philadelphia: Society for Industrial and Applied Mathematics.

Bazaraa, M., H. Sherali, and C. Shetty (1993). *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton: Princeton University Press.

Bernstein, D., T. L. Friesz, R. L. Tobin, and B. W. Wie (1993). A variational control formulation of the simultaneous route and departure-time equilibrium problem. *Proceedings of the International Symposium on Transportation and Traffic Theory*, 107–126.

Bertsekas, D. and R. Gallager (1992). *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall.

Datta-Chaudhuri, M. (1967). Optimum allocation of investments and transportation in a two-region economy. In K. Shell (Ed.), *Essays on the Theory of Optimal Economic Growth*, pp. 129–140. MIT Press.

Domazlicky, B. (1977). A note on the inclusion of transportation in models of the regional allocation of investment. *Journal of Regional Science 17*, 235–241.

Friesz, T., D. Bernstein, Z. Suo, and R. Tobin (2001). Dynamic network user equilibrium with state-dependent time lags. *Networks and Spatial Economics 1(3/4)*, 319–347.

Friesz, T. and N. Kydes (2003). The dynamic telecommunications flow routing problem. *Networks and Spatial Economics 4*(1), 55–73.

Friesz, T. and J. Luque (1987). Optimal regional growth models: multiple objectives, singular controls, and sufficiency conditions. *Journal of Regional Science 27*, 201–224.

Friesz, T. and R. Mookherjee (2006). Solving the dynamic network user equilibrium problem with state-dependent time shifts. *Transportation Research Part B 40*(3), 207–229.

Friesz, T., R. Mookherjee, and M. Rigdon (2004). Differential variational inequalities with state-dependent time shifts and applications to differential games. In *11th International Symposium on Dynamic Games and Applications, Tucson*.

Friesz, T. L., D. Bernstein, T. Smith, R. Tobin, and B. Wie (1993). A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research 41*(1), 80–91.

Friesz, T. L., M. A. Rigdon, and R. Mookherjee (2006). Differential variational inequalities and shipper dynamic oligopolistic network competition. *Transportation Research Part B 40*, 480–503.

Hotelling, H. (1978). A mathemathical theory of population. *Environment and Planning A 10*, 1223–1239.

Intriligator, M. (1964). Regional allocation of investment: comment. *Quarterly Journal of Economics 78*, 659–662.

Isaacs, R. (1965). *Differential Games*. New York: Dover.

McGill, J. and G. van Ryzin (1999). Revenue management: research overview and prospects. *Transportation Science 33*(2), 233–256.

Miller, T., T. L. Friesz, and R. L. Tobin (1996). *Equilibrium Facility Location on Networks*. New York: Springer-Verlag.

Mookherjee, R. and T. Friesz (2008). Pricing, allocation, and overbooking in dynamic service network competition when demand is uncertain. *Production and Operations Management 14*(4), 1–20.

Ohtsuki, Y. A. (1971). Regional allocation of public investment in an *n*-region economy. *Journal of Regional Science 11*, 225–233.

Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko (1962). *The Mathematical Theory of Optimal Processes*. New York: Interscience.

Puu, T. (1989). *Lecture Notes in Economics and Mathematical Systems*. New York: Springer-Verlag.

Puu, T. (1997). *Mathematical Location and Land Use Theory; An Introduction*. New York: Springer-Verlag.

Rahman, M. (1963). Regional allocation of investment: an aggregative study in the theory of development programming. *Quarterly Journal of Economics 77*, 26–39.

Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal 38*, 543–559.

Sakashita, N. (1967). Regional allocation of public investments. *Papers, Regional Science Association 19*, 161–182.

Sethi, S. P. and G. L. Thompson (1981). *Optimal Control theory: Applications to Management Science*. Boston: Martinus Nijhoff.

Tait, K. (1965). *Singular Problems in Optimal Control*. Ph. D. thesis, Harvard University, Cambridge, MA.

Chapter 2
Nonlinear Programming and Discrete-Time Optimal Control

The primary intent of this chapter is to introduce the reader to the theoretical foundations of nonlinear programming as well as the theoretical foundations of deterministic discrete-time optimal control. In fact, deterministic discrete-time optimal control problems, as we shall see, are actually nonlinear mathematical programs with a very particular type of structure. In a later chapter, we will also discover that deterministic continuous-time optimal control problems are specific instances of mathematical programs in topological vector spaces. Consequently, it is imperative for the student of optimal control to have a command of the foundations of nonlinear programming. Particularly important are the notions of local and global optimality in mathematical programming, the Kuhn-Tucker necessary conditions for optimality in nonlinear programming, and the role played by convexity in making necessary conditions sufficient. Readers already comfortable with finite-dimensional nonlinear programming may wish to go immediately to Section 2.9. We do caution, however, that subsequent chapters of this book assume substantial familiarity with finite-dimensional nonlinear programming, so that an overestimate of one's nonlinear programming knowledge can be very detrimental to ultimately obtaining a deep understanding of optimal control theory and differential games.

The following is an outline of the principal topics covered in this chapter:

Section 2.1: Nonlinear Program Defined. A formal definition of a finite-dimensional nonlinear mathematical program, with a single criterion and both equality and inequality constraints, is given.

Section 2.2: Other Types of Mathematical Programs. Definitions of linear, interger and mixed integer mathematical programs are provided.

Section 2.3: Necessary Conditions for an Unconstrained Minimum. We derive necessary conditions for a minimum of a twice continuously differentiable function when there are no constraints.

Section 2.4: Necessary Conditions for a Constrained Minimum. Relying on geometric reasoning, the Kuhn-Tucker conditions, as well as the notion of a constraint qualification, are introduced.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_2, © Springer Science+Business Media, LLC 2010

Section 2.5: Formal Derivation of the Kuhn-Tucker Conditions. A formal derivation of the Kuhn-Tucker necessary conditions, employing a conic definition of optimality and theorems of the alternative, is provided.

Section 2.6: Sufficiency, Convexity, and Uniqueness. We provide formal definitions of a convex set and a convex function. Then we show formally how those notions influence sufficiency and uniqueness of a global minimum.

Section 2.7: Generalized Convexity and Sufficiency. We extend the notion of convexity to include quasiconvexity and pseudoconvexity; we then show how these extensions may be used to state less restrictive conditions assuring optimality.

Section 2.8: Numerical and Graphical Examples. We provide numerical and graphical examples that illustrate the abstract optimality conditions introduced in previous sections of this chapter.

Section 2.9: Discrete-Time Optimal Control. We use the necessary conditions for nonlinear programs to derive the so-called minimum principle for discrete-time optimal control and associated necessary conditions.

2.1 Nonlinear Program Defined

We are presently interested in a type of optimization problem known as a finite-dimensional mathematical program, namely: find a vector $x \in \Re^n$ that satisfies

$$\left.\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h(x) = 0 \\ & g(x) \leq 0 \end{array}\right\} \tag{2.1}$$

where

$$\begin{aligned} x &= (x_1, \ldots, x_n)^T \in \Re^n \\ f(\cdot): &\ \Re^n \to \Re^1 \\ g(x) &= (g_1(x), \ldots, g_m(x))^T : \Re^n \to \Re^m \\ h(x) &= (h_1(x), \ldots, h_q(x))^T : \Re^n \to \Re^q \end{aligned}$$

We call the x_i for $i \in \{1, 2, \ldots, n\}$ decision variables, $f(x)$ the objective function, $h(x) = 0$ the equality constraints and $g(x) \leq 0$ the inequality constraints. Becuase the objective and constraint functions will in general be nonlinear, we shall consider (2.1) to be our canonical form of a nonlinear mathematical program (NLP). The *feasible region* for (2.1) is

$$X \equiv \{x : g(x) \leq 0,\ h(x) = 0\} \subset \Re^n \tag{2.2}$$

which allows us to state (2.1) in the form

$$\left.\begin{array}{l} \min\ f(x) \\ \text{s.t. } x \in X \end{array}\right\} \tag{2.3}$$

The pertinent definitions of optimality for NLP are:

Definition 2.1. *Global minimum. Suppose* $x^* \in X$ *and* $f(x^*) \leq f(x)$ *for all* $x \in X$. *Then* $f(x)$ *achieves a global minimum on* X *at* x^*, *and we say* x^* *is a global minimizer of* $f(x)$ *on* X.

Definition 2.2. *Local minimum. Suppose* $x^* \in X$ *and there exists an* $\epsilon > 0$ *such that* $f(x^*) \leq f(x)$ *for all* $x \in [N_\epsilon(x^*) \cap X]$, *where* $N_\epsilon(x^*)$ *is a ball of radius* $\epsilon > 0$ *centered at* x^*. *Then* $f(x)$ *achieves a local minimum on* X *at* x^*, *and we say* x^* *is a local minimizer of* $f(x)$.

In practice, we will often relax the formal terminology of Definition 2.1 and Definition 2.2 and refer to x^* as a global minimum or a local minimum, respectively.

2.2 Other Types of Mathematical Programs

We note that the general form of a continuous mathematical program (MP) may be specialized to create various types of mathematical programs that have been studied in depth. In particular, if the objective function and all constraint functions are linear, (2.1) is called a *linear program* (LP). In such cases, we normally add slack/surplus variables to the inequality constraints to convert them into equality constraints. That is, if we have the constraint

$$g_i(x) \leq 0 \tag{2.4}$$

we convert it into

$$g_i(x) + s_i = 0 \tag{2.5}$$

and solve for both x and s_i. The variable s_i is called a *slack variable* and obeys

$$s_i \geq 0 \tag{2.6}$$

If we have an inequality constraint of the form

$$g_j(x) \geq 0 \tag{2.7}$$

we convert it to the form

$$g_j(x) - s_j = 0 \tag{2.8}$$

where

$$s_j \geq 0 \tag{2.9}$$

is called a *surplus variable*. Thus, we take can convert any problem with inequality constraints into one that has only equality constraints and non-negativity restrictions. So without loss of generality, we take the canonical form of the linear programming problem to be

$$\begin{array}{ll} \min & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} & \sum_{j=1}^{n} a_{ij} x_j = b_i \quad i = 1, \ldots, m \\ & U_j \geq x_j \geq L_j \qquad j = 1, \ldots, n \\ & x \in \Re^n \end{array} . \tag{2.10}$$

where $n > m$. This problem can be re-stated further, using matrix and vector notation, as

$$\left.\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & U \geq x \geq L \\ & x \in \Re^n \end{array}\right\} \text{LP} \tag{2.11}$$

where $c \in \Re^n$, $b \in \Re^n$, and $A \in \Re^{m \times n}$.

If the objective function and/or some of the constraints are nonlinear, (2.1) is called a nonlinear program (NLP) and is written as:

$$\left.\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \;\; i = 1, \ldots, m \\ & h_i(x) = 0 \;\; i = 1, \ldots, q \\ & x \in \Re^n \end{array}\right\} \text{NLP} \tag{2.12}$$

If all of the elements of x are restricted to be a subset of the integers and I^n denotes the integer real numbers, the resulting program

$$\left.\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \;\; i = 1, \ldots, m \\ & h_i(x) = 0 \;\; i = 1, \ldots, q \\ & x \in I^n \end{array}\right\} \text{IP} \tag{2.13}$$

is called an integer program (IP). If there are two classes of variables, some that are continuous and some that are integer, as in

$$\left.\begin{array}{lll} \min & f(x,y) & \\ \text{s.t.} & g_i(x,y) \leq 0 & i = 1,\ldots,m \\ & h_i(x,y) = 0 & i = 1,\ldots,q \\ & x \in \Re^n \qquad y \in I^n & \end{array}\right\} \text{MIP}, \tag{2.14}$$

the problem is known as a mixed integer program (MIP).

2.3 Necessary Conditions for an Unconstrained Minimum

Necessary conditions for optimality in the mathematical program (2.1) are systems of equalities and inequalities that must hold at an optimal solution $x^* \in X$. Any such condition has the logical structure:

If x^* is optimal, then some property $\mathbf{P}(x^*)$ is true.

Necessary conditions play a central role in the analysis of most mathematical programming models and algorithms. Understanding them is also extremely important to understanding the theory of optimal control, even when considering problems in the infinite-dimensional vector spaces associated with continuous-time optimization. This is because the optimal control necessary condition known as the *minimum principle* requires solution of a finite-dimensional nonlinear program.

We begin our discussion of necessary conditions for mathematical programs by considering a special case of the general finite-dimensional mathematical program introduced in the previous section. In particular, we want to state and prove the following result for mathematical programs without constraints:

Theorem 2.1. *Necessary conditions for an unconstrained minimum. Suppose* $f : \Re^n \longrightarrow \Re^1$ *is twice continuously differentiable for all* $x \in \Re^n$. *Then necessary conditions for* $x^* \in \Re^n$ *to be a local or global minimum of* $\min f(x)$ *s.t.* $x \in \Re^n$ *are*

$$\nabla f\left(x^*\right) = 0 \tag{2.15}$$

$$\nabla^2 f\left(x^*\right) \equiv \left(\frac{\partial^2 f\left(x^*\right)}{\partial x_i \partial x_j}\right) \textit{ must be positive semidefinite} \tag{2.16}$$

That is, the gradient vanishes and the Hessian is positive semidefinite matrix at the minimum of interest.

Proof. Since $f(.)$ is twice continuously differentiable, we may make a Taylor series expansion in the vicinity of $x^* \in \Re^n$, a local minimum:

$$f(x) = f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*) + \|x - x^*\|^2 \mathcal{O}(x - x^*)$$

where $\mathcal{O}(x - x^*) \longrightarrow 0$ as $x \longrightarrow x^*$. If $\nabla f(x^*) \neq 0$, then by picking $x = x^* - \theta \nabla f(x^*)$ we can make $f(x) < f(x^*)$ for sufficiently small $\theta > 0$ and, thereby, directly contradict the fact that x^* is a local minimum. It follows that condition (2.15) is necessary, and we may write

$$f(x) = f(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*) + \|x - x^*\|^2 \mathcal{O}(x - x^*)$$

If the matrix $\nabla^2 f(x^*)$ is not positive semidefinite, there must exist a direction vector $d \in \Re^n$ such that $d \neq 0$ and $d^T \nabla^2 f(x^*) d < 0$. If we now choose $x = x^* + \theta d$, it is possible for sufficiently small $\theta > 0$ to realize $f(x) < f(x^*)$ in direct contradiction of the fact x^* is a local minimum. ■

2.4 Necessary Conditions for a Constrained Minimum

We comment that necessary conditions for constrained programs have the same logical structure as necessary conditions for unconstrained programs introduced in Section 8.4.4; namely:

If x^* is optimal, then some property $\mathbf{P}(x^*)$ is true.

For constrained programs, we will shortly find that $\mathbf{P}(x^*)$ is either the so-called Fritz John conditions or the Kuhn-Tucker conditions. We now turn to the task of providing an informal motivation of the Fritz John conditions, which are the pertinent necessary conditions for the case when no constraint qualification is imposed.

2.4.1 The Fritz John Conditions

The fundamental theorem on necessary conditions is:

Theorem 2.2. *Fritz John conditions. Let x^* be a (global or local) minimum of*

$$\begin{aligned} &\min\ f(x) \\ &s.t.\ x \in \mathcal{F} = \{x \in X_0 : g(x) \leq 0, h(x) = 0\} \subset \Re^n \end{aligned}$$

where X_0 is a nonempty open set, $g : \Re^n \longrightarrow \Re^m$, and $h : \Re^n \longrightarrow \Re^q$. Assume that $f(x)$, $g_i(x)$ for $i \in [1, m]$ and $h_i(x)$ for $i \in [1, q]$ have continuous first

derivatives everywhere on X. *Then there must exist multipliers* $\mu_0 \in \Re^1_+$, $\mu = (\mu_1, \ldots, \mu_m)^T \in \Re^m_+$, *and* $\lambda^* = (\lambda_i^*, \ldots; \lambda_q^*)^T \in \Re^q$ *such that*

$$\mu_0 \nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) = 0 \tag{2.17}$$

$$\mu_i g_i(x^*) = 0 \quad \forall\, i \in [1, m] \tag{2.18}$$

$$\mu_i \geq 0 \quad \forall\, i \in [0, m] \tag{2.19}$$

$$(\mu_0,\ \mu,\ \lambda) \neq 0 \in \Re^{m+q+1} \tag{2.20}$$

Conditions (2.17), (2.18), (2.19), and (2.20) together with $h(x) = 0$ and $g(x) \leq 0$ are called the Fritz John conditions. We will give a formal proof of their validity in Section 2.5.3. For now our focus is on how the Fritz John conditions are related to the Kuhn-Tucker conditions, which are the chief applied notion of a necessary condition for optimality in mathematical programming.

2.4.2 Geometry of the Kuhn-Tucker Conditions

Under certain regularity conditions called constraint qualifications, we may be certain that $\mu_0 \neq 0$. In that case, without loss of generality, we may take $\mu_0 = 1$. When $\mu_0 = 1$, the Fritz John conditions are called the Kuhn-Tucker conditions and (2.17) is called the Kuhn-Tucker identity. In either case, (2.18) and (2.19) together are called the complementary slackness conditions. Sometimes it is convenient to define the Lagrangean function:

$$L(x, \lambda, \mu_0, \mu) \equiv \mu_0 f(x) + \lambda^T h(x) + \mu^T g(x) \tag{2.21}$$

By virtue of this definition, identity (2.17) can be expressed as

$$\nabla_x L(x^*, \lambda, \mu_0, \mu) = 0 \tag{2.22}$$

At the same time (2.18) and (2.19) can be written as

$$\mu^T g(x^*) = 0 \tag{2.23}$$

$$\mu \geq 0 \tag{2.24}$$

Furthermore, we may give a geometrical motivation for the Kuhn-Tucker conditions by considering the following abstract problem with two decision variables and two inequality constraints:

$$\left.\begin{array}{ll} \min & f(x_1, x_2) \\ \text{s.t.} & g_1(x_1, x_2) \leq 0 \\ & g_2(x_1, x_2) \leq 0 \end{array}\right\} \tag{2.25}$$

The functions $f(.)$, $g_1(.)$, and $g_2(.)$ are assumed to be such that the following are true:

1. all functions are differentiable;
2. the feasible region $X \equiv \{(x_1, x_2) : g_1(x_1, x_2) \leq 0,\ g_2(x_1, x_2) \leq 0\}$ is a convex set;
3. all level sets $S_k \equiv \{(x_1, x_2) : f(x_1, x_2) \leq f_k\}$ are convex, where $f_k \in [\alpha, +\infty) \subset \Re^1_+$ is a constant and α is the unconstrained minimum of $f(x_1, x_2)$; and
4. the level curves

$$C_k = \left\{(x_1, x_2) : f(x_1, x_2) = f_k \in \Re^1\right\}$$

for the ordering

$$f_0 < f_1 < f_2 < \cdots < f_k$$

do not cross one another, and C_k is the locus of points for which the objective function has the constant value f_k.

Figure 2.1 is one realization of the above stipulations. Note that there is an uncountable number of level curves and level sets since f_k may be any real number from the interval $[\alpha, +\infty) \subset \Re^1_+$. In Figure 2.1, because the gradient of any function points in the direction of maximal increase of the function, we see there is a $\mu_1 \in \Re^1_{++}$ such that

$$\nabla f(x_1^*, x_2^*) = -\mu_1 \nabla g_1(x_1^*, x_2^*), \tag{2.26}$$

where (x_1^*, x_2^*) is the optimal solution formed by the tangency of $g_1(x_1^*, x_2^*) = 0$ with the level curve $f(x_1^*, x_2^*) = f_2$. Evidently, this observation leads directly to

$$\nabla f(x_1^*, x_2^*) + \mu_1 \nabla g_1(x_1^*, x_2^*) + \mu_2 \nabla g_2(x_1^*, x_2^*) = 0 \tag{2.27}$$

$$\mu_1 g_1(x_1^*, x_2^*) = 0 \tag{2.28}$$

$$\mu_2 g_2(x_1^*, x_2^*) = 0 \tag{2.29}$$

$$\mu_1, \mu_2 \geq 0, \tag{2.30}$$

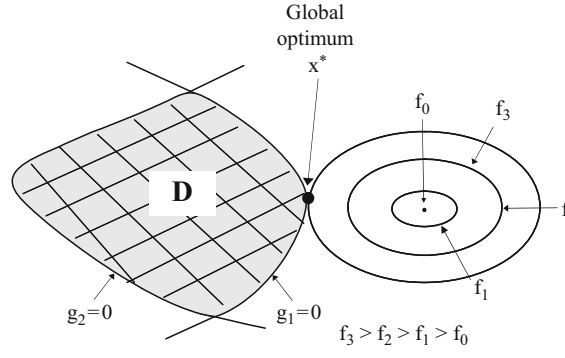


Fig. 2.1 Geometry of an optimal solution

Note that $g_1(x_1^*, x_2^*) = 0$ allows us to conclude that (2.28) holds even though $\mu_1 > 0$. Similarly, (2.26) implies that $\mu_2 = 0$, so (2.29) holds even though $g_2(x_1^*, x_2^*) \neq 0$. Clearly, the nonnegativity conditions (2.30) also hold. By inspection, (2.27), (2.28), (2.29), and (2.30) are the Kuhn-Tucker conditions (Fritz John conditions with $\mu_0 = 1$) for the mathematical program (2.25).

2.4.3 *The Lagrange Multiplier Rule*

We wish to give a statement of a particular instance of the Kuhn-Tucker theorem on necessary conditions for mathematical programming problems, together with some informal remarks about why that theorem holds when a constraint qualification is satisfied. Since our informal motivation of the Kuhn-Tucker conditions in the next section depends on the Lagrange multiplier rule (LMR) for mathematical programs with equality constraints, we must first state and motivate the LMR. To that end, take x and y to be scalars and $F(x, y)$ and $h(x, y)$ to be scalar functions. Consider the following mathematical program with two decision variables and a single equality constraint:

$$\left.\begin{array}{ll} \min & F(x, y) \\ \text{s.t.} & h(x, y) = 0 \end{array}\right\} \tag{2.31}$$

Assume that $h(x, y) = 0$ may be manipulated to find x in terms of y. That is, we know

$$x = H(y) \tag{2.32}$$

so that

$$F(x, y) = F[H(y), y] \equiv \Phi(y) \tag{2.33}$$

and (2.31) may be thought of as the one-dimensional unconstrained problem

$$\min_y \Phi(y) \tag{2.34}$$

which has the apparent necessary condition

$$\frac{d\Phi(y)}{dy} = 0 \tag{2.35}$$

By the chain rule we have the alternative form

$$\frac{d\Phi(y)}{dy} = \frac{\partial F(H, y)}{\partial y} + \frac{\partial F(H, y)}{\partial H}\frac{\partial H}{\partial y} = 0 \tag{2.36}$$

Applying the chain rule to the equality constraint $h(x, y) = 0$ leads to

$$dh(x, y) = \frac{\partial h}{\partial x}dx + \frac{\partial h}{\partial y}dy = 0 \tag{2.37}$$

from which we obtain

$$\frac{\partial x}{\partial y} = (-1)\frac{\partial h/\partial y}{\partial h/\partial x} \tag{2.38}$$

The necessary condition (2.36), with the help of (2.32) and (2.38), becomes

$$\begin{aligned}\frac{\partial F}{\partial y} + \frac{\partial F}{\partial x}\frac{\partial x}{\partial y} &= \frac{\partial F}{\partial y} + \frac{\partial F}{\partial x}(-1)\frac{\partial h/\partial y}{\partial h/\partial x} \\ &= \frac{\partial F}{\partial y} + (-1)\,\frac{\partial F/\partial x}{\partial h/\partial x}\frac{\partial h}{\partial y} \\ &= \frac{\partial F}{\partial y} + \lambda\frac{\partial h}{\partial y} = 0 \end{aligned} \tag{2.39}$$

where we have defined the Lagrange multiplier to be

$$\lambda = (-1)\,\frac{\partial F/\partial x}{\partial h/\partial x} \tag{2.40}$$

The LMR consists of (2.39) and (2.40), which we restate as

$$\frac{\partial F}{\partial x} + \lambda\frac{\partial h}{\partial x} = 0 \tag{2.41}$$

$$\frac{\partial F}{\partial y} + \lambda\frac{\partial h}{\partial y} = 0 \tag{2.42}$$

Recognizing that the generalization of (2.41) and (2.42) involves Jacobian matrices, we are not surprised to find that, for the equality constrained mathematical program

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \end{aligned}$$

where $x \in \Re^n$ and $h \in \Re^q$, the following result holds:

Theorem 2.3. *Lagrange multiplier rule. Let $x^* \in \Re^n$ be any local maximum or minimum of $f(x)$ subject to the constraints $h_i(x) = 0$ for $i \in [1, q]$, where $x \in \Re^n$ and $q < n$. If it is possible to choose a set of q variables for which the Jacobian*

$$J[h(x^*)] \equiv \begin{bmatrix} \dfrac{\partial h_1(x^*)}{\partial x_1} & \cdots & \dfrac{\partial h_1(x^*)}{\partial x_q} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_q(x^*)}{\partial x_1} & \cdots & \dfrac{\partial h_q(x^*)}{\partial x_q} \end{bmatrix} \tag{2.43}$$

has an inverse, then there exists a unique vector of Lagrange multipliers $\lambda = (\lambda_1, \ldots, \lambda_m)^T$ *satisfying*

$$\frac{\partial f(x^*)}{\partial x_j} + \sum_{i=1}^{q} \lambda_i \frac{\partial h_i(x^*)}{\partial x_j} = 0 \quad j \in [1, n] \tag{2.44}$$

The formal proof of this classical result is contained in most texts on advanced calculus. Note that (2.44) is a necessary condition for optimality.

2.4.4 Motivating the Kuhn-Tucker Conditions

We now wish, using the Lagrange multiplier rule, to establish that the Kuhn-Tucker conditions are valid when an appropriate constraint qualification holds. In fact we wish to consider the following result:

Theorem 2.4. *Kuhn-Tucker conditions. Let* $x^* \in X$ *be a local minimum of*

$$\begin{aligned} &\min \quad f(x) \\ &\text{s.t.} \quad x \in X = \{x \in X_0 : g(x) \leq 0, h(x) = 0\} \subset \Re^n \end{aligned}$$

where X_0 *is a nonempty open set. Assume that* $f(x)$, $g_i(x)$ *for* $i \in [1, m]$ *and* $h_i(x)$ *for* $i \in [1, q]$ *have continuous first derivatives everywhere on* X *and that a constraint qualification holds. Then there must exist multipliers* $\mu = (\mu_1, \ldots, \mu_m)^T \in \Re^q$ *and* $\lambda^* = (\lambda_i^*, \ldots, \lambda_q^*)^T \in \Re^m$ *such that*

$$\nabla f(x^*) + \textstyle\sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) = 0 \tag{2.45}$$

$$\mu_i g_i(x^*) = 0 \quad \forall\, i \in [1, m] \tag{2.46}$$

$$\mu_i \geq 0 \quad \forall\, i \in [1, m] \tag{2.47}$$

Expression (2.45) is the Kuhn-Tucker identity and conditions (2.46) and (2.47), as we have indicated previously, are together referred to as the complementary slackness conditions. Do not fail to note that the Kuhn-Tucker conditions are necessary conditions. A solution of the Kuhn-Tucker conditions, without further information, is only a candidate optimal solution, sometimes referred to as a "Kuhn-Tucker point." In fact, it is possible for a particular Kuhn-Tucker point not to be an optimal solution.

We may informally motivate Theorem 2.4 using the Lagrange multiplier rule. This is done by first positing the existence of variables s_i, unrestricted in sign, for $i \in [1, m]$ such that

$$g_i(x^*) + (s_i)^2 = 0 \quad \forall\, i \in [1, m] \tag{2.48}$$

so that the mathematical program (2.1) may be viewed as one with only equality constraints, namely

$$\left.\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h(x) = 0 \\ & g(x) + \text{diag}\,(s) \cdot s = 0 \end{array}\right\} \tag{2.49}$$

where $s \in \Re^m$ and

$$\text{diag}\,(s) \equiv \begin{pmatrix} s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{m-1} & 0 \\ 0 & 0 & \cdots & 0 & s_m \end{pmatrix} \tag{2.50}$$

To form the necessary conditions for this mathematical program, we first construct the Lagrangian

$$\begin{aligned} L\,(x,s,\lambda,\mu) &= f(x) + \lambda^T h(x) + \mu^T\,[g(x) + \text{diag}\,(s) \cdot s] \\ &= f(x) + \sum_{i=1}^{q} \lambda_i h_i\,(x) + \sum_{i=1}^{m} \mu_i \left[g_i\,(x) + s_i^2\right] \end{aligned} \tag{2.51}$$

and then state, using the LMR, the first-order conditions

$$\frac{\partial L\,(x,s,\lambda,\mu)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{q} \lambda_j \frac{\partial h_j\,(x)}{\partial x_i} + \sum_{j=1}^{m} \mu_j \frac{\partial g_j\,(x)}{\partial x_i} = 0 \qquad i \in [1,n] \tag{2.52}$$

$$\frac{\partial L\,(x,s,\lambda,\mu)}{\partial s_i} = 2\mu_i s_i = 0 \quad i \in [1,m] \tag{2.53}$$

Result (2.52) is of course the Kuhn-Tucker identity (2.45). Note further that both sides of (2.53) may be multiplied by $-s_i$ to obtain the equivalent conditions

$$\mu_i \left(-s_i^2\right) = 0 \quad i \in [1,m] \tag{2.54}$$

which can be restated using (2.48) as

$$\mu_i g_i\,(x) = 0 \quad i \in [1,m] \tag{2.55}$$

Conditions (2.55) are of course the complementary slackness conditions (2.46).

It remains for us to establish that the inequality constraint multipliers μ_i for $i \in [1,m]$ are nonnegative. To that end, we imagine a perturbation of the inequality constraints by the vector

$$\varepsilon = \begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_m \end{pmatrix}^T \in \Re_{++}^m,$$

so that the inequality constraints become

$$g(x) + \text{diag}\,(s) \cdot s = \varepsilon$$

or

$$g_i\,(x) + s_i^2 - \varepsilon_i = 0 \quad i \in [1, m] \tag{2.56}$$

There is an optimal solution for each vector of perturbations, which we call $x\,(\varepsilon)$ where $x^* = x\,(0)$ is the unperturbed optimal solution. As a consequence there is an optimal objective function value

$$Z\,(\varepsilon) \equiv f\left[x^*\,(\varepsilon)\right] \tag{2.57}$$

for each $x^*\,(\varepsilon)$. We note that

$$\frac{\partial Z\,(\varepsilon)}{\partial \varepsilon_i} = \sum_{j=1}^{n} \frac{\partial f\,(x)}{\partial x_j} \frac{\partial x_j\,(\varepsilon)}{\partial \varepsilon_i} \tag{2.58}$$

by the chain rule. Similarly for $k \in [1, m]$

$$\frac{\partial g_k\,(x)}{\partial \varepsilon_i} = \frac{\partial \left[\varepsilon_k - s_k^2\right]}{\partial \varepsilon_i} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \tag{2.59}$$

and for $k \in [1, q]$

$$\frac{\partial h_k\,(x)}{\partial \varepsilon_i} = \sum_{j=1}^{n} \frac{\partial h_k\,(x)}{\partial x_j} \frac{\partial x_j\,(\varepsilon)}{\partial \varepsilon_i} \tag{2.60}$$

Furthermore, we may define

$$\Phi_i \equiv \frac{\partial Z\,(\varepsilon)}{\partial \varepsilon_i} + \sum_{k=1}^{q} \lambda_k \frac{\partial h_k\,(x)}{\partial \varepsilon_i} + \sum_{k=1}^{m} \mu_k \frac{\partial g_k\,(x)}{\partial \varepsilon_i} \tag{2.61}$$

and note that

$$\Phi_i = \frac{\partial Z\,(\varepsilon)}{\partial \varepsilon_i} + \mu_i \tag{2.62}$$

With the help of (2.58), (2.59), and (2.60), we have

$$\begin{aligned} \Phi_i &= \sum_{j=1}^{n} \frac{\partial f\,(x)}{\partial x_j} \frac{\partial x_j\,(\varepsilon)}{\partial \varepsilon_i} + \sum_{k=1}^{q} \lambda_k \sum_{j=1}^{n} \frac{\partial h_k\,(x)}{\partial x_j} \frac{\partial x_j\,(\varepsilon)}{\partial \varepsilon_i} + \sum_{k=1}^{m} \mu_k \sum_{j=1}^{n} \frac{\partial g_k\,(x)}{\partial x_j} \frac{\partial x_j\,(\varepsilon)}{\partial \varepsilon_i} \\ &= \sum_{j=1}^{n} \left[\frac{\partial f\,(x)}{\partial x_j} + \sum_{k=1}^{q} \lambda_k \frac{\partial h_k\,(x)}{\partial x_j} + \sum_{k=1}^{m} \mu_k \frac{\partial g_k\,(x)}{\partial x_j} \right] \frac{\partial x_j\,(\varepsilon)}{\partial \varepsilon_i} = 0 \end{aligned} \tag{2.63}$$

by virtue of the Kuhn-Tucker identity (2.52). From (2.62) and (2.63) it is immediate that

$$\mu_i = (-1)\frac{\partial Z(\varepsilon)}{\partial \varepsilon_i} \quad i \in [1, m] \tag{2.64}$$

We now note that, when the unconstrained minimum of $f(x)$ is external to the feasible region

$$X(\varepsilon) = \{x : g(x) \leq \varepsilon, h(x) = 0\},$$

increasing ε_i can never increase, and may potentially lower, the objective function for all $i \in [1, m]$; that is

$$\frac{\partial Z(\varepsilon)}{\partial \varepsilon_i} \leq 0 \quad i \in [1, m] \tag{2.65}$$

From (2.64) and (2.65) we have the desired result

$$\mu_i \geq 0 \quad \forall i \in [1, m] \tag{2.66}$$

ensuring that the multipliers for inequality constraints are nonnegative.

2.5 Formal Derivation of the Kuhn-Tucker Conditions

We are interested in formally proving that, under the linear independence constraint qualification and some other basic assumptions, the Kuhn-Tucker identity and the complementary slackness conditions form, together with the original mathematical program's constraints, a valid set of necessary conditions. For finite-dimensional mathematical programs, the only type we consider in this chapter, such a demonstration is facilitated by Gordon's lemma, which is in effect a corollary of Farkas' lemma of classical analysis. The problem structure needed to apply Gordon's lemma can be most readily created by expressing the notion of optimality in terms of cones and separating hyperplanes. Throughout this section we consider the mathematical program

$$\min f(x) \quad \text{s.t.} \quad x \in \mathcal{F} \tag{2.67}$$

where, depending on context, either $\mathcal{F}$ is a general set or

$$\mathcal{F} \equiv \{x \in X_0 : g(x) \leq 0\} \subset \Re^n \tag{2.68}$$

and

$$f : \Re^n \to \Re^1 \tag{2.69}$$

$$g : \Re^n \to \Re^m \tag{2.70}$$

where X_0 is a nonempty open set in $\Re^n$. Note that we presently consider only inequality constraints, as any equality constraint

$$h_k(x) = 0$$

may be stated as two inequality constraints

$$h_k(x) \leq 0$$
$$-1 \cdot h_k(x) \leq 0$$

2.5.1 Cones and Optimality

A cone is a set obeying the following definition:

Definition 2.3. *Cone. A set C in $\Re^n$ is a cone with vertex zero if $x \in C$ implies that $\theta x \in C$ for all $\theta \in \Re^1_+$.*

Now consider the following definitions:

Definition 2.4. *Cone of feasible directions. For the mathematical program (2.67), provided $\mathcal{F}$ is not empty, the cone of feasible directions at $x \in X$ is*

$$D_0(x) = \{d \neq 0 : x + \theta d \in \mathcal{F} \quad \forall \theta \in (0, \delta) \text{ and some } \delta > 0\}$$

Definition 2.5. *Feasible direction. Every nonzero vector $d \in D_0$ is called a feasible direction at $x \in X$ for the mathematical program (2.67).*

Definition 2.6. *Cone of improving directions. For the mathematical program (10.1), if f is differentiable at $x \in \mathcal{F}$, the cone of improving directions at $x \in \mathcal{F}$ is*

$$F_0(x) = \{d : [\nabla f(x)]^T \cdot d < 0\}$$

Definition 2.7. *Feasible direction of descent. Every vector $d \in F_0 \bigcap D_0$ is called a feasible direction of descent at $x \in \mathcal{F}$ for the mathematical program (2.67).*

Definition 2.8. *Cone of interior directions. For the mathematical program (10.1), if g_i is differentiable at $x \in X$ for all $i \in I(x)$, where*

$$I(x) = \{i : g_i(x) = 0\},$$

then, the cone of interior directions at $x \in \mathcal{F}$ is

$$G_0(x) = \{d : [\nabla g_i(x)]^T \cdot d < 0 \quad \forall\}$$

Note that in Definition 2.4 and Definition 2.6, if $\mathcal{F}$ is a convex set, we may set $\delta = 1$ and refer only to $\theta \in [0, 1]$, as will become clear in the next section after we define the notion of a convex set. Furthermore, the definitions immediately above allow one to characterize an optimal solution of (2.67) as a circumstance for which the intersection of the cone of feasible directions and the cone of improving directions is empty. This has great intuitive appeal for it says that there are no feasible directions that allow the objective to be improved. In fact, the following result obtains:

Theorem 2.5. *Optimality in terms of the cones of feasible and improving directions. Consider the mathematical program*

$$\min f(x) \quad s.t. \quad x \in \mathcal{F} \tag{2.71}$$

where $f : \Re^n \longrightarrow \Re^1$, $\mathcal{F} \subseteq \Re^n$ *and* $\mathcal{F}$ *is nonempty. Suppose also that* f *is differentiable at the local minimum* $x^* \in \mathcal{F}$ *of (2.71). Then at* x^* *the intersection of the cone of feasible directions* D_0 *and the cone of improving directions* F_0 *is empty:*

$$F_0(x^*) \cap D_0(x^*) = \emptyset$$

That is, at the local solution $x^* \in \mathcal{F}$*, no improving direction is also a feasible direction.*

Proof. The result is intuitive. For a formal proof see Bazaraa et al. (1993). ■

Theorem 2.6. *Optimality in terms of the cones of interior and improving directions. Let* $x^* \in \mathcal{F}$ *be a local minimum of the mathematical program*

$$min\, f(x) \quad s.t. \quad x \in \mathcal{F} = \{x \in X_0 : g(x) \leq 0\} \subset \Re^n \tag{2.72}$$

where X_0 *is a nonempty open set in* $\Re^n$, $f : \Re^n \longrightarrow \Re^1$, *and* $g : \Re^n \longrightarrow \Re^m$ *are differentiable at* x^*, *while the* g_i *for* $i \in I$ *are continuous at* x^*. *The cone of improving directions and the cone of interior directions satisfy*

$$F_0(x^*) \cap G_0(x^*) = \emptyset$$

Proof. This result is also intuitive. For a formal proof see Bazaraa et al. (1993). ■

2.5.2 Theorems of the Alternative

Farka's Lemma is a specific example of a so-called *theorem of the alternative*. Such theorems provide information on whether a given linear system has a solution when a related linear system has or fails to have a solution. Farkas' lemma has the following statement:

Lemma 2.1. *Farkas' lemma. Let* A *be an* $m \times n$ *matrix of real numbers and* $c \in \Re^n$. *Then exactly one of the following systems has a solution: System 1:* $Ax \leq 0$ *and* $c^T x > 0$ *for some* $x \in \Re^n$*; or System 2:* $A^T y = c$ *and* $y \geq 0$ *for some* $y \in \Re^m$.

Proof. Farkas' lemma is proven in most advanced texts on nonlinear programming. See, for example, Mangasarian (1969). ■

Corollary 2.1. *Gordon's corollary. Let* A *be an* $m \times n$ *matrix of real numbers. Then exactly one of the following systems has a solution: System 1:* $Ax < 0$ *for some* $x \in \Re^n$*; or System 2:* $A^T y = 0$ *and* $y \geq 0$ *for some* $y \in \Re^m$.

Proof. See Mangasarian (1969). ■

2.5.3 The Fritz John Conditions Again

By using Corollary 2.1 it is quite easy to establish the Fritz John conditions introduced previously and restated here without equality constraints:

Theorem 2.7. *The Fritz John conditions. Let $x^* \in \mathcal{F}$ be a minimum of*

$$\begin{aligned} &\min \quad f(x) \\ &s.t. \quad x \in \mathcal{F} = \{x \in X_0 : g(x) \leq 0\} \end{aligned}$$

where X_0 is a nonempty open set in $\Re^n$ and $g : \Re^n \longrightarrow \Re^m$. Assume that $f(x)$ and $g_i(x)$ for $i \in [1, m]$ have continuous first derivatives everywhere on $\mathcal{F}$. Then there must exist multipliers $\mu_0 \in \Re^1_+$ and $\mu = (\mu_1, \ldots, \mu_m)^T \in \Re^m_+$ such that

$$\mu_0 \nabla f(x^*) + \textstyle\sum_{i=1}^m \mu_i \nabla g_i(x^*) = 0 \tag{2.73}$$

$$\mu_i g_i(x^*) = 0 \quad \forall\, i \in [1, m] \tag{2.74}$$

$$\mu_i \geq 0 \quad \forall\, i \in [1, m] \tag{2.75}$$

$$(\mu_0, \mu) \neq 0 \in \Re^{m+1} \tag{2.76}$$

Proof. Since $x^* \in \mathcal{F}$ solves the mathematical program of interest, we know from Theorem 2.6 that $F_0(x^*) \cap G_0(x^*) = \emptyset$; that is, there is no vector d satisfying

$$\left[\nabla f(x^*)\right]^T \cdot d < 0 \tag{2.77}$$

$$\left[\nabla g_i(x^*)\right]^T \cdot d < 0 \quad i \in I(x^*) \tag{2.78}$$

where $I(x^*)$ is the set of indices of constraints binding at x^*. Without loss of generality, we may consecutively number the binding constraints from 1 to $|I(x^*)|$ and define

$$A = \begin{pmatrix} [\nabla f(x^*)]^T & 0 & 0 & \cdots & 0 \\ 0 & [\nabla g_1(x^*)]^T & 0 & \cdots & 0 \\ 0 & 0 & [\nabla g_2(x^*)]^T & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \left[\nabla g_{|I(x^*)|}(x^*)\right]^T \end{pmatrix}$$

As a consequence we may state (2.77) and (2.78) as

$$A \begin{pmatrix} d \\ d \\ \vdots \\ d \end{pmatrix} < 0 \tag{2.79}$$

According to Corollary 2.1, since (2.79) cannot occur, there exists

$$y = \begin{pmatrix} \mu_0 \\ \mu_i : i \in I(x^*) \end{pmatrix} \geq 0$$

such that

$$A^T y = A^T \begin{pmatrix} \mu_0 \\ \mu_i : i \in I(x^*) \end{pmatrix} = 0 \tag{2.80}$$

Expression (2.80) yields

$$\mu_0 \nabla f\left(x^*\right) + \sum_{i=1}^{|I(x^*)|} \mu_i \nabla g_i(x^*) = 0 \tag{2.81}$$

We are free to introduce the additional multipliers

$$\mu_i = 0 \quad i = |I(x^*)| + 1, \ldots, m \tag{2.82}$$

which assure that the complementary slackness conditions (2.74) and (2.75) hold for all multipliers. As a consequence of (2.81) and (2.82), we have (2.73), thereby completing the proof. ■

2.5.4 The Kuhn-Tucker Conditions Again

With the apparatus developed so far, we wish to prove the following restatement of Theorem 2.4 in terms of the linear independence constraint qualification:

Theorem 2.8. *Kuhn-Tucker conditions. Let $x^* \in \mathcal{F}$ be a local minimum of*

$$\begin{aligned} &min \ f(x) \\ &s.t. \ x \in \mathcal{F} = \{x \in X_0 : g(x) \leq 0, h(x) = 0\} \end{aligned}$$

where X_0 is a nonempty open set in $\Re^n$. Assume that $f(x)$, $g_i(x)$ for $i \in [1, m]$ and $h_i(x)$ for $i \in [1, q]$ have continuous first derivatives everywhere on $\mathcal{F}$ and that the gradients of binding constraint functions are linearly independent. Then there

must exist multipliers $\mu = (\mu_1, \ldots, \mu_m)^T \in \Re^m$ *and* $\lambda = (\lambda_i, \ldots, \lambda_q)^T \in \Re^q$ *such that*

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) = 0 \tag{2.83}$$

$$\mu_i g_i(x^*) = 0 \quad \forall\, i \in [1, m] \tag{2.84}$$

$$\mu_i \geq 0 \quad \forall\, i \in [1, m] \tag{2.85}$$

Proof. Recall that a constraint qualification is a condition that guarantees the multiplier μ_0 of the Fritz John conditions is non-zero. We again use the notation

$$I\left(x^*\right) = \left\{i : g_i\left(x^*\right) = 0\right\}, \tag{2.86}$$

for the set of subscripts corresponding to binding inequality constraints. Note also that by their very nature equality constraints are always binding. Linear independence of the gradients of binding constraints means that only zero multipliers

$$\mu_i = 0 \quad \forall\, i \in I\left(x^*\right) \tag{2.87}$$

$$\lambda_i = 0 \quad \forall\, i \in [1, q] \tag{2.88}$$

allow the identity

$$\sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) = 0, \tag{2.89}$$

to hold. We are free to set the multipliers for nonbinding constraints to zero; that is

$$g_i(x^*) < 0 \Longrightarrow \mu_i = 0 \quad \forall\, i \notin I\left(x^*\right)$$

which assures (2.84) and (2.85) hold for $i \in [1, m]$. Consequently, linear independence of the gradients of binding constraints actually means that there are no nonzero multipliers assuring

$$\sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) = 0 \tag{2.90}$$

That is, either all $\lambda_i = 0$ and all $\mu_i = 0$ or

$$\sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) \neq 0 \tag{2.91}$$

In the latter case, the Fritz John identity

$$\mu_0 \nabla f\left(x^*\right) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x^*) = 0 \tag{2.92}$$

immediately forces

$$\mu_0 \neq 0 \tag{2.93}$$

unless $\nabla f\ (x^*) = 0 \in \Re^n$; in this latter case (2.90) must hold and so we may still enforce (2.93) without contradiction or loss of generality. ■

2.6 Sufficiency, Convexity, and Uniqueness

Sufficient conditions for optimality in a mathematical program are conditions that, if satisfied, ensure optimality. Any such condition has the logical structure:

If property $\mathbf{P}(x^*)$ is true, then x^* is optimal.

It turns out that convexity, a notion that requires careful definition, provides useful sufficient conditions that are relatively easy to check in practice. In particular, we will define a convex mathematical program to be a mathematical program with a convex objective function (when minimizing) and a convex feasible region, and we will show that the Kuhn-Tucker conditions are not only necessary but also sufficient for global optimality in such programs.

2.6.1 Quadratic Forms

A key concept, useful for establishing convexity of functions, is that of a quadratic form, formally defined as follows:

Definition 2.9. *Quadratic form. A quadratic form is a scalar-valued function defined for all $x \in \Re^n$ that takes on the following form:*

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{2.94}$$

where each a_{ij} is a real number.

Note that any quadratic form may be expressed in matrix notation as

$$Q(x) = x^T A x \tag{2.95}$$

where $A = (a_{ij})$ is an $n \times n$ matrix. It is well known that for any given quadratic form there is a symmetric matrix S that allows one to to re-express that quadratic form as

$$Q(x) = x^T S x \tag{2.96}$$

where the elements of $S = (s_{ij})$ are given by $s_{ij} = s_{ji} = (a_{ij} + a_{ji})/2$. Because of this symmetry property, we may assume, without loss of generality, that every quadratic form is already expressed in terms of a symmetric matrix. That is, whenever we encounter a quadratic form such as (2.95) or (2.96), the underlying matrix generating that form may be taken to be symmetric if so doing assists our analysis.

A quadratic form may exhibit various properties, two of which are the subject of the following definition:

Definition 2.10. *Positive definiteness. The quadratic form $Q(x) = x^T S x$ is positive definite on $\Omega \subseteq \Re^n$ if $Q(x) > 0$ for all $x \in \Omega$ and $x \neq 0$. The quadratic form $Q(x) = x^T S x$ is positive semidefinite on $\Omega \subseteq \Re^n$ if $Q(x) \geq 0$ for all $x \in \Omega$.*

Analogous definitions may be made for negative definite and negative semidefinite quadratic forms. An important lemma concerning quadratic forms, which we state without proof, is the following:

Lemma 2.2. *Properties of positive definite matrix. Let the symmetric $n \times n$ matrix S be positive (negative) definite. Then*

1. *The inverse S^{-1} exists;*
2. *S^{-1} is positive (negative) definite; and*
3. *$A^T S A$ is positive (negative) semidefinite for any $m \times n$ matrix A.*

In addition, we will need the following lemma, which we also state without proof:

Lemma 2.3. *Nonnegativity of principal minirs. A quadratic form $Q(x) = x^T S x$, where S is the associated symmetric matrix, is positive semidefinite if and only if it may be ordered so that s_{11} is positive and the following determinants of the principal minors are all nonnegative:*

$$\begin{vmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{vmatrix} \geq 0, \quad \begin{vmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{vmatrix} \geq 0, \ \ldots, |S| \geq 0$$

2.6.2 Concave and Convex Functions

This section contains several definitions, lemmas, and theorems related to convex functions and convex sets that we need to fully understand the notion of sufficiency. First, consider the following four definitions:

Definition 2.11. *Convex set. A set $X \subseteq \Re^n$ is convex if for any two vectors $x^1, x^2 \in X$ and any scalar $\lambda \in [0, 1]$ the vector*

$$x = \lambda x^1 + (1-\lambda)x^2 \tag{2.97}$$

also lies in X.

Definition 2.12. *Strictly convex set. A set* $X \subseteq \Re^n$ *is strictly convex if for any two vectors* x^1 *and* x^2 *in* X *and any scalar* $\lambda \in (0,1)$ *the point*

$$x = \lambda x^1 + (1-\lambda)x^2 \tag{2.98}$$

lies in the interior of X.

Definition 2.13. *Convex function. A scalar function* $f(x)$ *is a convex function defined over a convex set* $X \subseteq \Re^n$ *if for any two vectors* $x^1, x^2 \in X$

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2) \quad \forall\ \lambda \in [0,1] \tag{2.99}$$

Definition 2.14. *Strictly convex function. In the above,* $f(x)$ *is strictly convex if the inequality is a strict inequality* $(<)$ *for all* $\lambda \in (0,1)$.

Note that concave and strictly concave functions are defined by reversing the inequalities in the preceding definitions.

We are now ready to state the following theorems:

Theorem 2.9. *Sum of convex functions. The sum of any two convex functions is convex.*

Theorem 2.10. *Convexity of linear functions. Any linear function is both convex and concave.*

Theorem 2.11. *Convexity of quadratic form. For any positive semidefinite and symmetric matrix* S, *the quadratic form* $Q(x) = x^T S x$ *is a convex function over all of* $\Re^n$.

The proofs of the preceding results are straightforward and are left to the reader. Another important result is the following that relates convex level sets and convex functions:

Theorem 2.12. *Level sets of convex funtion. If* $f(x)$ *is a (strictly) convex function over* $\Re^n$, *then the set of points*

$$S \equiv \{x : f(x) \leq b\}, \tag{2.100}$$

where b *is any real number, is a (strictly) convex set.*

Proof. The definition of convexity tells us that

$$\begin{aligned} f(\lambda x^1 + (1-\lambda)x^2) &\leq \lambda f(x^1) + (1-\lambda) f(x^2) \\ &\leq \lambda b + (1-\lambda)b = b \end{aligned}$$

A strict version of this inequality is obtained for strictly convex functions, thereby completing the proof. ■

We will also need the following lemma:

Lemma 2.4. *Intersection of convex sets. The intersection of any two convex sets is itself a convex set.*

Proof. Take x^1 and x^2 within the intersection $X^1 \cap X^2$, where X^1 and X^2 are convex sets. Join these points by a line segment. That line segment and all the points on it are both in X^1 and X^2. ■

It is now trivial to establish the following result:

Theorem 2.13. *Convex feasible region. The feasible region X of the mathematical program (2.3) is a convex set if the following two conditions are met:*

1. *the equality constraint functions $h_i(x) : \Re^n \longrightarrow \Re^1$ for $i \in [1, q]$ are all linear on X; and*
2. *the inequality constraint functions $g_i(x) : \Re^n \longrightarrow \Re^1$ for $i \in [1, m]$ are all convex on X.*

Proof. For the given, the sets

$$S_h = \{x : h(x) = 0\}$$
$$S_g = \{x : g(x) \leq 0\}$$

are convex. The feasible region X obeys

$$X = S_h \cap S_g$$

Hence, X is convex, since the intersection of two convex sets is a convex set. ■

Now we are ready to deal with the following key result:

Theorem 2.14. *Global minimum of a convex program. If the function $f(x) : \Re^n \longrightarrow \Re^1$ is defined and convex on the closed convex set $X \subseteq \Re^n$, then any constrained local minimum of $f(x)$ for $x \in X$ is a global minimum on X. Similarly, if $f(x)$ is concave on the closed convex set X, then any constrained local maximum of $f(x)$ for $x \in X$ is a global maximum on X.*

Proof. Suppose $x^0 \in X$ is a constrained local minimum but not a global minimum, so that there exists some $x^* \in X$ such that $f(x^*) < f(x^0)$. Then for any $\lambda \in [0, 1]$ the convexity of $f(x)$ tells us that

$$\begin{aligned} f(\lambda x^* + (1-\lambda)x^0) &\leq \lambda f(x^*) + (1-\lambda) f(x^0) \\ &< \lambda f(x^0) + (1-\lambda) f(x^0) = f(x^0) \end{aligned} \tag{2.101}$$

Now, consider a straight line segment from x^0 to x^* which must lie entirely in X (by convexity). For any small positive δ (a scalar), there exists $\lambda > 0$ such that

$$x = \lambda x^* + (1 - \lambda)x^0 \tag{2.102}$$

lies in X at a distance δ away from x^0. However, we have already shown in (2.101) that

$$f(x) < f(x^0) \tag{2.103}$$

Since δ may be infinitesimally small, x^0 cannot be a local minimum. Hence, we have a contradiction. ■

Another important result is the following:

Theorem 2.15. *Tangent line property of a convex function. Let $f(x)$ have continuous first partial derivatives. Then $f(x)$ is convex over the convex region $X \subseteq \Re^n$ if and only if*

$$f(x) \geq f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) \tag{2.104}$$

for any two vectors x^* and x in X. Moreover, $f(x)$ is concave over the convex region $X \subseteq \Re^n$ if and only if

$$f(x) \leq f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) \tag{2.105}$$

for any two vectors x^* and x in X.

This result may be proven by taking a Taylor series expansion of $f(x)$ about the point x^* and arguing that the second order and higher terms sum to a positive number. Theorem 2.15 expresses the geometric property that a tangent to a convex function will underestimates that function. Still another related result is:

Theorem 2.16. *Convexity and positive semidefiniteness of the Hessian. Let $f(x)$ have continuous second partial derivatives. Then $f(x)$ is convex (concave) over some the region $X \subseteq \Re^n$ if and only if its Hessian matrix*

$$H(x) \equiv \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \tag{2.106}$$

is positive (negative) semidefinite.

Proof. We give the proof for concave functions, although the case of convex functions is completely analogous.

(i) [negative semidefiniteness $\Longrightarrow$ concavity] First note that the Hessian H is symmetric by its very nature. We may make a second-order Taylor series expansion of $f(x)$ about a point $x^* \in X$ to obtain

$$f(x) = f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) + \frac{1}{2}(x - x^*)^T H[x^* + \theta(x - x^*)](x - x^*) \tag{2.107}$$

for some $\theta \in (0, 1)$. Because X is convex we know that the point

$$x^* + \theta(x - x^*) = \theta x + (1 - \theta)x^* , \tag{2.108}$$

a convex combination of x and x^*, must lie within X. Now suppose that H is negative definite or negative semidefinite throughout X, so that the last term on the righthand side of the Taylor expansion is clearly negative or zero. We get

$$f(x) \leq f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) \tag{2.109}$$

It follows from the previous theorem that $f(x)$ is concave.

(ii) [concavity $\Longrightarrow$ negative semidefiniteness] Now assume $f(x)$ is concave throughout X but that the Hessian matrix H is not negative semidefinite at some point $x^* \in X$. Then, of course, there will exist a vector y such that

$$y^T H(x^*)y > 0 \tag{2.110}$$

Now define $x^0 = x^* + y$ and rewrite this last inequality as

$$(x^0 - x^*)^T H(x^*)(x^0 - x^*) > 0 \tag{2.111}$$

Consider another point $x = x^* + \beta(x^0 - x^*)$ where β is a real positive number, so that

$$(x^0 - x^*) = \frac{1}{\beta}(x - x^*) \tag{2.112}$$

It follows that for any such β

$$(x - x^*)^T H(x^*)(x - x^*) > 0 \tag{2.113}$$

Since H is continuous, we may choose x so close to x^* that

$$(x - x^*)^T H[x^* + \theta(x - x^*)](x - x^*) > 0 \tag{2.114}$$

for all $\theta \in [0, 1]$. By hypothesis $f(x)$ is concave over $\Re$ so that

$$f(x) \leq f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) \tag{2.115}$$

holds, together with the Taylor series expansion (2.107). Subtracting (2.115) from (2.107) gives

$$0 \geq \frac{1}{2}(x - x^*)^T H[x^* + \theta(x - x^*)](x - x^*) \tag{2.116}$$

for some $\theta \in (0, 1)$. This contradicts (2.114). ■

Note this last theorem cannot be strengthened to say a function is strictly convex if and only if its Hessian is positive definite. Examples may be given of functions that are strictly convex and whose Hessians are not positive definite. However, one can establish that positive definiteness of the Hessian does imply strict convexity by employing some of the arguments from the preceding proof.

Furthermore, the manner of construction of the preceding proofs leads directly to the following corollary:

Corollary 2.2. *Solution set convex. If the constrained global minimum of* $f(x)$ *for* $x \in X \subset \Re^n$ *is* α *when* $f(x) : \Re^n \longrightarrow \Re^1$ *is convex on* X*, a convex set, then the set*

$$\Psi = \{x : x \in X \subset \Re^n, \quad f(x) \leq \alpha\} \tag{2.117}$$

is the set of all solutions and is itself convex.

We now turn our attention to the question of additional regularity conditions that will assure that the set Ψ is a singleton. In fact, we will prove the following theorem:

Theorem 2.17. *Unique global minimum. Let* $f(\cdot)$ *be a strictly convex function defined on a convex set* $X \subset \Re^n$*. If* $f(\cdot)$ *attains its global minimum on* X*, it is attained at a unique point of* X*.*

Proof. Suppose there are two global minima: $x^1 \in X$ and $x^2 \in X$. Let $f(x^1) = f(x^2) = \alpha$. Then, by the previous corollary the set Ψ is a convex set and is the set of all solutions. Therefore

$$x^1, \ x^2, \ x^3 \in \Psi \tag{2.118}$$

where $x^3 = \lambda x^1 + (1 - \lambda)x^2$, and

$$\alpha = f(x^3) = f(\lambda x^1 + (1 - \lambda)x^2) < \lambda f(x^1) + (1 - \lambda) f(x^2) = \alpha.$$

This is a contradiction and therefore there cannot be two global minima. ■

2.6.3 Kuhn-Tucker Conditions Sufficient

The most significant implication of imposing regularity conditions based on convexity is that they make the Kuhn-Tucker conditions sufficient as well as necessary for global optimality. In fact, we may state and prove the following:

Theorem 2.18. *Kuhn-Tucker conditions sufficient for convex programs. Let*

$$\begin{aligned} f &: X \subset \Re^n \longrightarrow \Re^n \\ g &: X \subset \Re^n \longrightarrow \Re^m \\ h &: X \subset \Re^n \longrightarrow \Re^q \end{aligned}$$

be real-valued, differentiable functions. Suppose X_0 is an open convex set, while f is convex, the g_i are convex for $i \in [1,m]$, and the h_i are linear for $i \in [1,q]$. Take x^* to be a feasible solution of the mathematical program

$$\left.\begin{array}{l} min\ f(x) \\ s.t.\ h_i(x) = 0\ (\lambda_i)\ \ i \in [1,q] \\ \quad\ g_i(x) \leq 0\ (\mu_i)\ \ i \in [1,m] \\ \quad\ x \in X_0 \end{array}\right\} \tag{2.119}$$

If there exist multipliers $\mu^* \in \Re^m$ and $\lambda^* \in \Re^q$ satisfying the Kuhn-Tucker conditions

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i^* \nabla h_i(x^*) = 0$$

$$\mu_i^* g_i(x^*) = 0 \quad \mu_i^* \geq 0 \quad i \in [1,m] \quad ,$$

then x^* is a global minimum.

Proof. To simplify the exposition, we shall assume only constraints that are inequalities; this is possible since any linear equality constraint

$$h_k(x) = 0$$

for $k \in [1,m]$ may be restated as two convex inequality constraints in standard form:

$$\begin{aligned} h_k(x) &\leq 0 \\ -h_k(x) &\leq 0 \end{aligned}$$

and absorbed into the definition of $g(x)$. The Kuhn-Tucker identity is then

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla g_i(x^*) = 0 \tag{2.120}$$

Post multiplying (2.120) by $(x - x^*)$ gives

$$\left[\nabla f(x^*)\right]^T (x - x^*) + \sum_i \mu_i^* \left[\nabla g_i(x^*)\right]^T (x - x^*) = 0 \tag{2.121}$$

where x^* is a solution of the Kuhn-Tucker conditions and

$$x, x^* \in X = \{x \in X_0 : g(x) \leq 0, h(x) = 0\}$$

We know that for a convex, differentiable function

$$g(x) \geq g(x^*) + \left[\nabla g(x^*)\right]^T (x - x^*) \tag{2.122}$$

From (2.121) and (2.122), we have

$$\begin{aligned} \left[\nabla f(x^*)\right]^T (x - x^*) &= -\sum \mu_i^* \left[\nabla g_i(x^*)\right]^T (x - x^*) \\ &\geq \sum_i \mu_i^* [g_i(x^*) - g_i(x)] \\ &= \mu_i^* [-g_i(x)] \geq 0 \end{aligned} \tag{2.123}$$

because $\mu_i^* g_i(x^*) = 0$, $\mu_i^* \geq 0$ and $g_i(x) \leq 0$. Hence

$$\left[\nabla f(x^*)\right]^T (x - x^*) \geq 0 \tag{2.124}$$

Because $f(x)$ is convex

$$f(x) \geq f(x^*) + \left[\nabla f(x^*)\right]^T (x - x^*) \tag{2.125}$$

Hence, from (2.124) and (2.125) we get

$$f(x) - f(x^*) \geq \left[\nabla f(x^*)\right]^T (x - x^*) \geq 0 \tag{2.126}$$

That is

$$f(x) \geq f(x^*) \ ,$$

which establishes that any solution of the Kuhn-Tucker conditions is a global minimum for the given. ■

Note that this theorem can be changed to one in which the objective function is strictly convex, thereby assuring that any corresponding solution of the Kuhn-Tucker conditions is an unique global minimum. Its given may also be relaxed if certain results from the theory of generalized convexity are employed.

2.7 Generalized Convexity and Sufficiency

There are certain generalizations of the notion of convexity that allow the sufficiency conditions introduced above to be somewhat weakened. We begin to explore the notion of more general types of convexity by introducing the following definition of a quasiconvex function:

Definition 2.15. *Quasiconvex function. The function $f : X \longrightarrow \Re^n$ is quasiconvex on the set $X \subset \Re^n$ if*

$$f\left(\lambda_1 x^1 + \lambda_2 x^2\right) \leq \max\left[f\left(x^1\right), f\left(x^2\right)\right]$$

for every $x^1, x^2 \in X$ and every $(\lambda_1, \lambda_2) \in \left\{(\lambda_1, \lambda_2) \in \Re^2_+ : \lambda_1 + \lambda_2 = 1\right\}$.

We next introduce the notion of a pseudoconvex function:

Definition 2.16. *Pseudoconvex function. The function $f : X \longrightarrow \Re^n$, differentiable on the open convex set $X \subset \Re^n$, is pseudoconvex on X if*

$$\left(x^1 - x^2\right)^T \nabla f\left(x^2\right) \geq 0$$

implies that

$$f(x^1) \geq f(x^2)$$

for every $x^1, x^2 \in X$.

Pseudoconcavity of f occurs of course when $-f$ is pseudoconvex. Furthermore, we shall say a function is pseudolinear (quasilinear) if it is both pseudoconvex (quasiconvex) and pseudoconcave (quasiconcave).

The notions of generalized convexity we have given allow the following theorem to be stated and proven:

Theorem 2.19. *Kuhn-Tucker conditions sufficient for generalized convex programs. Let*

$$\begin{aligned} f &: X \subset \Re^n \longrightarrow \Re^n \\ h &: X \subset \Re^n \longrightarrow \Re^m \\ g &: X \subset \Re^n \longrightarrow \Re^q \end{aligned}$$

be real-valued, differentiable functions. Suppose X_0 is an open convex set, while f is pseudoconvex, the g_i are quasiconvex for $i \in [1, m]$, and the h_i are quasilinear for $i \in [1, q]$. Take x^* to be a feasible solution of the mathematical program

$$\begin{array}{lll} min & f(x) & \\ s.t. & h_i(x) = 0 \quad (\eta_i) & i \in [1, q] \\ & g_i(x) \leq 0 \quad (\lambda_i) & i \in [1, m] \\ & x \in X_0 & \end{array}$$

If there exist multipliers $\mu^* \in \Re^m$ and $\lambda^* \in \Re^q$ satisfying the Kuhn-Tucker conditions

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla g_i(x^*) + \sum_{i=1}^{q} \lambda_i^* \nabla h_i(x^*) = 0$$

$$\mu_i^* g_i(x^*) = 0 \quad \mu_i^* \geq 0 \quad i \in [1, m],$$

then x^* is a global minimum.

Proof. The proof is left as an exercise for the reader. ■

We close this section by noting that if in addition to the given of Theorem 2.19 an appropriate notion of strict pseudoconvexity is introduced for the objective function f, then the Kuhn-Tucker conditions become sufficient for a unique global minimizer.

2.8 Numerical and Graphical Examples

In this section we provide several numerical and graphical examples meant to test and refine the reader's knowledge of the material on nonlinear programming presented above. We will need the notions of a level curve C_k and a level set S_k of the objective function $f(x)$ of a mathematical program:

$$C_k = \{x : f(x) = f_k\} \tag{2.127}$$

$$S_k = \{x : f(x) \leq f_k\} \tag{2.128}$$

where f_k signifies a numerical value of the objective function of interest. Solving any mathematical program graphically involves four steps:

1. Draw the feasible region.
2. Draw level curves of the objective function**.**
3. Choose the optimal level curve by selecting, from the points of tangency of level curves and constraint boundaries, the fesible point or points giving the best objective function value.
4. Identify the optimal solution as the point of tangency between the optimal level curve and the feasible region

2.8.1 LP Graphical Solution

Consider the following linear program:

$$\max f(x, y) = x + y$$

subject to

$$3x + 2y \leq 6 \tag{2.129}$$

$$\frac{1}{2}x + y \leq 2 \tag{2.130}$$

For the present example the optimal solution is, by inspection of Figure 2.2, the point

$$x^* = \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix} \tag{2.131}$$

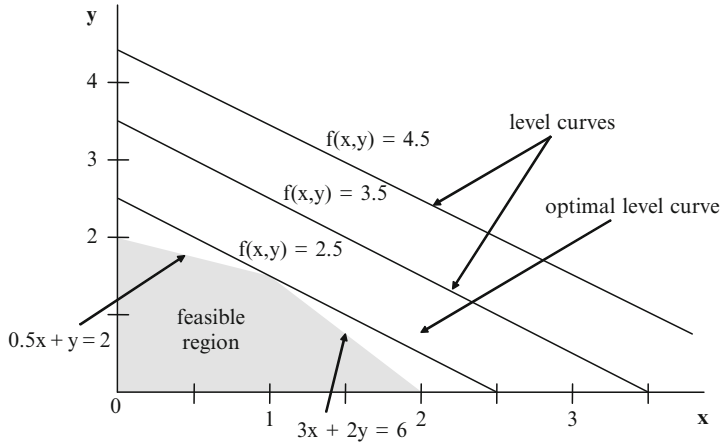


Fig. 2.2 LP graphical solution

One can easily verify the Kuhn-Tucker conditions hold at this point. To do so, it is helpful to restate the problem as follows:

$$\min f(x, y) = -x - y \tag{2.132}$$

$$g_1(x, y) = 3x + 2y - 6 \leq 0 \tag{2.133}$$

$$g_2(x, y) = \frac{1}{2}x + y - 2 \leq 0 \tag{2.134}$$

We note that

$$\nabla f(x, y) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \tag{2.135}$$

$$\nabla g_1(x, y) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \tag{2.136}$$

$$\nabla g_2(x, y) = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} \tag{2.137}$$

The Kuhn-Tucker identity is

$$\nabla f\ (x_1, x_2) + \lambda_1 \nabla g_1\ (x_1, x_2) + \lambda_2 \nabla g_2\ (x_1, x_2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.138}$$

That is

$$\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.139}$$

The complementary slackness conditions are

$$\lambda_1 g_1(x_1, x_2) = 0 \quad \lambda_1 \geq 0 \tag{2.140}$$

$$\lambda_2 g_2(x_1, x_2) = 0 \quad \lambda_2 \geq 0 \tag{2.141}$$

Note that

$$I = \{i : g_i\left(1, \frac{3}{2}\right) = 0\} = \{1, 2\} \tag{2.142}$$

and we must find multipliers that obey

$$\lambda_1, \lambda_2 \geq 0 \tag{2.143}$$

It is easy to solve the above system and show

$$\lambda_1 = \frac{1}{4} > 0, \lambda_2 = \frac{1}{2} > 0 \tag{2.144}$$

Hence x^* satisfies the Kuhn-Tucker conditions. Because the problem is a linear program, it is a convex program. Therefore, the Kuhn-Tucker conditions are not only necessary but also sufficient, making x^* a global solution.

2.8.2 NLP Graphical Example

Consider the following nonlinear program

$$\min \ f(x_1, x_2) = (x_1 - 5)^2 + (x_2 - 6)^2 \tag{2.145}$$

subject to

$$g_1(x_1, x_2) = \frac{1}{2}x_1 + x_2 - 3 \leq 0 \tag{2.146}$$

$$g_2(x_1, x_2) = x_1 - 2 \leq 0 \tag{2.147}$$

By inspection of Figure 2.3, the point (2, 2) is the globally optimal solution with a corresponding objective function value of 25. Note that

$$\nabla f(2, 2) = \begin{pmatrix} -8 \\ -6 \end{pmatrix} \tag{2.148}$$

$$\nabla g_1(2, 2) = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} \tag{2.149}$$

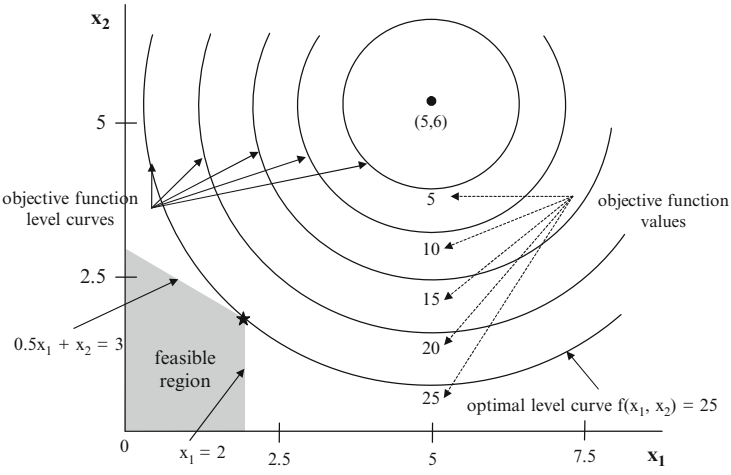


Fig. 2.3 NLP graphical solution

$$\nabla g_2(2,2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.150}$$

The Kuhn-Tucker identity is

$$\begin{pmatrix} -8 \\ -6 \end{pmatrix} + \lambda_1 \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.151}$$

The complementary slackness conditions are

$$\lambda_1 g_1(x_1, x_2) = 0 \quad \lambda_1 \geq 0 \tag{2.152}$$

$$\lambda_2 g_2(x_1, x_2) = 0 \quad \lambda_2 \geq 0 \tag{2.153}$$

and

$$I = \{i : g_i(1,2) = 0\} = \{1,2\} \Longrightarrow \lambda_1, \lambda_2 \geq 0 \tag{2.154}$$

Solving the above linear system (2.151) yields multipliers of the correct sign:

$$\lambda_1 = 6 > 0 \tag{2.155}$$

$$\lambda_2 = 5 > 0 \tag{2.156}$$

Consequently, the Kuhn-Tucker conditions are satisfied. Because the program is convex with a strictly convex objective function, we know that the Kuhn-Tucker conditions are both necessary and sufficient for an unique global optimum. So, even without further analysis, we know $(2, 2)$ is the unique global optimum.

2.8.3 Nonconvex, Nongraphical Example

Consider the nonlinear program

$$\min f\left(x_1, x_2\right) = -x_1 + 0x_2 \tag{2.157}$$

subject to

$$g_1\left(x_1, x_2\right) = \left(x_1\right)^2 + \left(x_2\right)^2 - 2 \leq 0 \tag{2.158}$$

$$g_2\left(x_1, x_2\right) = x_1 - \left(x_2\right)^2 \leq 0 \tag{2.159}$$

Note that the feasible region of this mathematical program is not convex; hence, we will have to enumerate all the combinations of binding and nonbinding constraints in order to solve it using the Kuhn-Tucker conditions alone. We begin by observing that

$$\nabla f\left(x_1, x_2\right) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \tag{2.160}$$

$$\nabla g_1\left(x_1, x_2\right) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \tag{2.161}$$

$$\nabla g_2\left(x_1, x_2\right) = \begin{pmatrix} 1 \\ -2x_2 \end{pmatrix} \tag{2.162}$$

The Kuhn-Tucker identity is

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.163}$$

from which we obtain the equations

$$kti1 : -1 + 2\lambda_1 x_1 + \lambda_2 = 0 \tag{2.164}$$

$$kti2 : \left(\lambda_1 - \lambda_2\right) x_2 = 0 \tag{2.165}$$

The complementary slackness conditions are

$$csc1 : \lambda_1 g_1\left(x_1, x_2\right) = 0 \quad \lambda_1 \geq 0 \tag{2.166}$$

$$csc2 : \lambda_2 g_2\left(x_1, x_2\right) = 0 \quad \lambda_2 \geq 0 \tag{2.167}$$

Because there are $N = 2$ inequality constraints, there are $2^N = 2^2 = 4$ possible cases of binding and nonbinding constraints:

Case	g_1	g_2
I	< 0	< 0
II	< 0	$= 0$
III	$= 0$	< 0
IV	$= 0$	$= 0$

(2.168)

It is convenient to use the following symbols and operators for analyzing each of the four cases:

Symbol/Operator	Meaning
$\oplus$	consider two statements
$\Longrightarrow$	the implication of such a consideration
$\Lightarrow\!\!\!

Case IV: $\left[g_1 = (x_1)^2 + (x_2)^2 - 2 = 0\right] \oplus \left[g_2 = x_1 - (x_2)^2 = 0\right] \Longrightarrow$

$[x_1 = 1, x_2 = \pm 1] \oplus [kti1, kti2] \Longrightarrow [-1 + 2\lambda_1 + \lambda_2 = 0, \lambda_1 - \lambda_2 = 0] \Longrightarrow$

$[\lambda_1 = 1/3 > 0, \lambda_2 = 1/3 > 0] \Longrightarrow csc1$ and $csc2$ satisfied $\Longrightarrow$

$\left[x^B = (1,1)^T, x^C = (1,-1)^T \text{ are valid Kuhn-Tucker points}\right]$.

The global optimum is found by noting

$$\begin{aligned} f\left(x^A\right) &= 0 \\ f\left(x^B\right) = f\left(x^C\right) &= -1 < f\left(x^A\right) \end{aligned} \tag{2.170}$$

which means x^B, x^C are alternative global minimizers. Note also that x^A is not a local minimizer.

2.8.4 A Convex, Nongraphical Example

Let us now consider the mathematical program

$$\min f(x_1, x_2) = 0x_1 - x_2 \tag{2.171}$$

subject to

$$g_1(x_1, x_2) = (x_1)^2 + (x_2)^2 - 2 \leq 0 \tag{2.172}$$

$$g_2(x_1, x_2) = -x_1 + x_2 \leq 0 \tag{2.173}$$

Note that this problem is a convex mathematical program since the objective function is linear and the inequality constraint functions are convex. We know the Kuhn-Tucker conditions will be both necessary and sufficient for a nonunique global minimum. This means that we need only find one case of binding and nonbinding constraints that leads to nonnegative inequality constraint multipliers in order to solve (2.171), (2.172), and (2.173) to global optimality. We begin by observing that

$$\nabla f(x_1, x_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{2.174}$$

$$\nabla g_1(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \tag{2.175}$$

$$\nabla g_2(x_1, x_2) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{2.176}$$

The Kuhn-Tucker identity is

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.177}$$

from which we obtain the equations

$$kti1 : 2\lambda_1 x_1 - \lambda_2 = 0 \tag{2.178}$$

$$kti2 : -1 + 2\lambda_1 x_2 + \lambda_2 = 0 \tag{2.179}$$

The complementary slackness conditions are

$$csc1 : \lambda_1 g_1 (x_1, x_2) = 0 \quad \lambda_1 \geq 0 \tag{2.180}$$

$$csc2 : \lambda_2 g_2 (x_1, x_2) = 0 \quad \lambda_2 \geq 0 \tag{2.181}$$

Since the present mathematical program has two constraints, the table (2.168) still applies. Let us posit that both constraints are binding, so that the following analysis applies:

$\boxed{\text{Case IV}}$: $\left[g_1 = (x_1)^2 + (x_2)^2 - 2 = 0\right] \oplus [g_2 = -x_1 + x_2 = 0] \Longrightarrow$

$\left[x_1^* = x_2^* = 1\right] \oplus [kti1, kti2] \Longrightarrow [2\lambda_1 - \lambda_2 = 0, -1 + 2\lambda_1 + \lambda_2 = 0] \Longrightarrow$

$\left[\lambda_1 = \frac{1}{4} > 0, \lambda_2 = \frac{1}{2} > 0\right] \Longrightarrow [csc1 \text{ and } csc2 \text{ are satisfied}] \Longrightarrow$

$\left[x^* = \left(x_1^*, x_2^*\right)^T = (1, 1)^T \text{ is a global minimizer}\right]$

However, since the objective function is only convex and not strictly convex, we cannot ascertain without analyzing the three remaining cases whether this global minimizer is unique. The reader may verify that the other three cases lead to contradictions, and thereby determine that $x^* = (1, 1)^T$ is a unique global solution.

2.9 Discrete-Time Optimal Control

We are now ready to formulate a fairly general version of the discrete-time optimal control problem. Because time is treated discretely, we avoid in this initial foray into optimal control theory the complications and nuances of infinite-dimensional vector spaces. In particular, we will show that the discrete-time optimal control problem can be restated as a nonlinear mathematical program in standard form. We then show that application of the Kuhn-Tucker conditions leads us directly to a discrete version

of Pontryagin's minimum (maximum) principle and the other necessary conditions of discrete-time optimal control.

The *equations of motion*, also called the *dynamics*, that we consider take the form of the following difference equations:

$$x_{t+1} = x_t + f_t(x_t, u_t) \quad t = 0, 1, \ldots, q-1 \tag{2.182}$$

where t is a discrete time index (a nonnegative integer) and q is the number of time steps that constitute our *planning or analysis horizon*. Note further that $x_t \in \Re^n$ and $u_t \in \Re^r$ are vectors, as is $f_t : \Re^n \times \Re^r \longrightarrow \Re^n$. We refer to the x_t as state variables and the u_t as control variables. We assume f_t is continuously differentiable on $\Re^n \times \Re^r$. The initial and terminal conditions for these dynamics are taken, respectively, to be

$$\Phi_0(x_0) = 0 \tag{2.183}$$

$$\Phi_q(x_q) = 0 \tag{2.184}$$

where $\Phi_0 : \Re^n \longrightarrow \Re^{m_0}$ and $\Phi_q : \Re^n \longrightarrow \Re^{m_q}$, while Φ_0 and Φ_q are both $\mathcal{C}^1$ on $\Re^n$. The *control constraints* are stated in abstract form as

$$u_t \in \mathcal{U}_t \equiv \{u : g_t(u_t) \leq 0\} \subseteq \Re^r \quad t = 0, 1, \ldots, q-1 \tag{2.185}$$

where $g_t : \Re^r \longrightarrow \Re^s$ and g_t is $\mathcal{C}^1$ on $\Re^r$. As stressed in our development of the Kuhn-Tucker conditions, there is no loss of generality arising from the fact that we have only explicitly considered inequality constraints on the controls, as any equality constraint may be represented by two appropriately defined inequalities. The final piece of the discrete-time optimal control problem is its cost function defined by

$$J = \Psi(x_q) + \sum_{t=0}^{q-1} F_t(x_t, u_t) \tag{2.186}$$

where $\Psi : \Re^n \longrightarrow \Re^1$ is $\mathcal{C}^1$ on $\Re^n$, while $F_t : \Re^n \times \Re^r \longrightarrow \Re^1$ is $\mathcal{C}^1$ on $\Re^n \times \Re^r$. We assume that J is meant to be minimized.

Assembling the individual pieces presented above, we have the following canonical form of the discrete-time optimal control problem:

$$\min J = \Psi(x_q) + \sum_{t=0}^{q-1} F_t(x_t, u_t) \tag{2.187}$$

subject to

$$x_{t+1} = x_t + f_t(x_t, u_t) \quad t = 0, 1, ..., q-1 \tag{2.188}$$

$$u_t \in \mathcal{U}_t = \{u : g_t(u_t) \leq 0\} \subseteq \Re^r \quad t = 0, 1, \ldots, q-1 \tag{2.189}$$

$$\Phi_0(x_0) = 0 \quad \Phi_q(x_q) = 0 \tag{2.190}$$

Note that we have included no constraints involving the state variables.

2.9.1 Necessary Conditions

It should be apparent that the discrete-time optimal control problem given by (2.187), (2.188), (2.189), and (2.190) is a finite-dimensional nonlinear mathematical program. Let us put it in the following form:

$$\min Z(x, u) = \Psi(x_q) + \sum_{t=0}^{q-1} F_t(x_t, u_t) \tag{2.191}$$

subject to

$$h_t(x_t, x_{t+1}) = -x_{t+1} + x_t + f_t(x_t, u_t) = 0 \quad (\tau_{t+1}) \qquad t = 0, 1, ..., q-1 \tag{2.192}$$

$$g_t(u_t) \leq 0 \quad (\lambda_t) \quad t = 0, 1, ..., q-1 \tag{2.193}$$

$$\Phi_0(x_0) = 0 \quad (\rho_0) \tag{2.194}$$

$$\Phi_q(x_q) = 0 \quad (\rho_q) \tag{2.195}$$

where for convenience we employ the following notation

$$x = \begin{pmatrix} x_0 \\ \vdots \\ x_q \end{pmatrix} \in \Re^{n(q+1)} \tag{2.196}$$

$$u = \begin{pmatrix} u_0 \\ \vdots \\ u_{q-1} \end{pmatrix} \in \Re^{rq} \tag{2.197}$$

for the vectors of decision variables for our mathematical program; as we have mentioned, in the parlance of optimal control theory, these vectors are vectors of state variables and control variables, respectively.

We assume that a relevant constraint qualification is in force so that the Kuhn-Tucker conditions for the mathematical program (2.191), (2.192), (2.193), (2.194), and (2.195) are a valid characterization of optimality. The names of

multipliers for the constraints of (2.192), (2.193), (2.194), and (2.195) are indicated in parentheses next to each constraint. We state the Kuhn-Tucker conditions by first forming the Lagrangean; that is, we price out all constraints and adjoin them to the original objective function to obtain

$$\mathcal{L}(x,u,\rho,\tau) = \Psi\left(x_q\right) + \sum_{t=0}^{q-1} F_t\left(x_t,u_t\right) + \rho_0\Phi_0\left(x_0\right) + \rho_q\Phi_q\left(x_q\right) + \sum_{t=0}^{q-1}\tau_{t+1}^T\left(-x_{t+1}+x_t+f_t\left(x_t,u_t\right)\right) + \sum_{t=0}^{q-1}\lambda_t^T g_t\left(u_t\right)$$

where the symbol T denotes the transpose operation and

$$\rho = \begin{pmatrix}\rho_0\\ \rho_q\end{pmatrix} \in \Re^{m_0+m_q}$$

$$\tau = \begin{pmatrix}\tau_1\\ .\\ .\\ .\\ \tau_q\end{pmatrix} \in \Re^{nq}$$

are vectors of dual variables (ρ) and *adjoint variables*[1] (τ), respectively.

The Kuhn-Tucker identity is, of course, obtained by setting the partial derivatives of $\mathcal{L}(x,u,\rho,\tau)$ equal to zero; let us begin with the following:

$$\nabla_x\mathcal{L}(x,u,\rho,\tau) = 0 \tag{2.198}$$

It follows that

$$\frac{\partial\mathcal{L}}{\partial x_0} = \frac{\partial F_0}{\partial x_0} + \rho_0\frac{\partial\Phi_0}{\partial x_0} + \tau_1^T + \tau_1^T\frac{\partial f_0}{\partial x_0} = 0 \tag{2.199}$$

$$\frac{\partial\mathcal{L}}{\partial x_t} = \frac{\partial F_t}{\partial x_t} - \tau_t^T + \tau_{t+1}^T + \tau_{t+1}^T\frac{\partial f_t}{\partial x_t} = 0 \tag{2.200}$$

If we agree to define

$$\tau_0 = -\left[\frac{\partial\Phi_0}{\partial x_0}\right]^T\rho_0 \tag{2.201}$$

[1] These discrete-time adjoint variables are clearly mathematical programming dual variables; in optimal control theory, we refer to them as adjoint variables by tradition.

then (2.199) and (2.200) can be written as

$$\tau_0 = \tau_1 + \left[\frac{\partial f_0}{\partial x_0}\right]^T \tau_1 + \nabla_{x_0} F_0 \tag{2.202}$$

$$\tau_t = \tau_{t+1} + \left[\frac{\partial f_t}{\partial x_t}\right]^T \tau_{t+1} + \nabla_{x_t} F_t \quad t = 1, \ldots, q-1 \tag{2.203}$$

We note that (2.203) and (2.204) have the same form as one another, so they may be conveniently represented by the single statement

$$\tau_t = \tau_{t+1} + \left[\frac{\partial f_t}{\partial x_t}\right]^T \tau_{t+1} + \nabla_{x_t} F_t \quad t = 0, \ldots, q-1 \tag{2.204}$$

Next note that

$$\frac{\partial \mathcal{L}}{\partial x_q} = \frac{\partial \Psi}{\partial x_q} + \rho_q^T \frac{\partial \Phi_q}{\partial x_q} - \tau_q^T = 0 \tag{2.205}$$

which can be rewritten as

$$\tau_q = \nabla_{x_q} \Psi + \left[\frac{\partial \Phi_q}{\partial x_q}\right]^T \rho_q \tag{2.206}$$

The remaining partial derivatives of interest are those of the Lagrangean with respect to the control variables, which are set to zero:

$$\nabla_u \mathcal{L}(x, u, \rho, \tau) = 0 \tag{2.207}$$

It follows that

$$\frac{\partial \mathcal{L}}{\partial u_t} = \lambda_t^T \frac{\partial g_t}{\partial u_t} + \tau_{t+1}^T \frac{\partial f_t}{\partial u_t} + \frac{\partial F_t}{\partial u_t} = 0 \quad t = 0, \ldots, q-1 \tag{2.208}$$

which can be rewritten as

$$\nabla_{u_t} \left[\sum_{t=0}^{q-1} F_t(x_t, u_t) + \sum_{t=0}^{q-1} \tau_{t+1}^T \left(-x_{t+1} + x_t + f_t(x_t, u_t)\right) + \sum_{t=0}^{q-1} \lambda_t^T g_t(u_t) \right] = 0 \tag{2.209}$$

The final conditions for us to mention are

$$\lambda_t^T g_t = 0 \quad \lambda_t \geq 0 \quad t = 1, \ldots, q-1 \tag{2.210}$$

which are recognized as the complementary slackness conditions associated with the control inequality constraints and their multipliers.

In deriving the equations and inequalities of this section that express the necessary conditions, the arguments of all functions and their derivatives have been purposely omitted in order to simplify the notation. The complete set of necessary conditions for the discrete-time optimal control problem consist of the original problem constraints together with the conditions we have derived. That is to say, the necessary conditions are

$$\begin{aligned} \text{equations of motion} &: (2.182) \\ \text{initial conditions} &: (2.183) \\ \text{terminal conditions} &: (2.184) \\ \text{control constraints} &: (2.185) \\ \text{adjoint equations} &: (2.202) \\ \text{transversality conditions} &: (2.206) \\ \text{stationarity conditions for the controls} &: (2.209) \end{aligned}$$

Note that these conditions constitute a so-called *two-point boundary-value problem*.

2.9.2 The Minimum Principle

In this section we wish to manipulate the necessary conditions for the discrete-time optimal control problem developed from application of the Kuhn-Tucker conditions into the traditional form used to study and analyze optimal control problems; in the process we will articulate Pontryagin's *minimum principle*. The mathematics of this section are essentially algebra and some simple differentiation; the substantive aspect of the discrete-time optimal control problem analysis has already been completed in the previous section. However, the success of modern optimal control theory is in no small part due to the elegant, concise statement of the necessary conditions that we are about to give (and which is usually attributed to Pontryagin and his colleagues); packaging is important!

We begin the task of reformulating the necessary conditions by defining the *Hamiltonian*:

$$H_t(x_t, \tau_{t+1}, u_t) \equiv F_t(x_t, u_t) + \tau_{t+1}^T f_t(x_t, u_t) \quad t = 0, \ldots, q-1 \tag{2.211}$$

where $x_t \in \Re^n$ will be called the *state variable vector* while τ_t and u_t were named, in Section 2.9.1, the *adjoint vector* and *control vector*, respectively; furthermore $H_t : \Re^n \times \Re^n \times \Re^r \longrightarrow \Re^1$. It is immediate that the equations of motion may be stated as

$$x_{t+1} - x_t = \nabla_{\tau_{t+1}} H_t(x_t, \tau_{t+1}, u_t) \quad t = 0, \ldots, q-1 \tag{2.212}$$

and the adjoint equations as

$$\tau_{t+1} - \tau_t = -\nabla_{x_t} H_t\left(x_t, \tau_{t+1}, u_t\right) \quad t = 0, \ldots, q-1 \tag{2.213}$$

Results (2.212) and (2.213) are completely analogous to Hamilton's equations of classical mechanics that describe conservative Newtonian systems in terms of generalized coordinates (position and momentum). For this reason, these equations are sometimes still called Hamilton's equations, although there is no implication that (2.212) and (2.213) carry with them any of the assumptions or implications of classical mechanics.

We may also, in light of the definition of the Hamiltonian (2.211), restate the stationarity conditions for the optimal controls as

$$\nabla_{u_t}\left[H_t\left(x_t, \tau_{t+1}, u_t\right) + \sum_{t=0}^{q-1} \lambda_t^T g_t\left(u_t\right)\right] = 0 \quad t = 0, \ldots, q-1 \tag{2.214}$$

$$\lambda_t^T g_t\left(u_t\right) = 0 \quad \lambda_t \geq 0 \quad t = 0, \ldots, q-1 \tag{2.215}$$

The system (2.214) and (2.215) is immediately recognized as the necessary conditions for statically minimizing the Hamiltonian with respect to the controls under the assumption that all other variables are held fixed. We restate this observation as

$$H_t\left(x_t, \tau_{t+1}, u_t\right) \leq H_t\left(x_t, \tau_{t+1}, u\right) \quad \forall u \in \mathcal{U}_t \quad t = 0, \ldots, q-1 \tag{2.216}$$

Expression (2.216) is Pontryagin's minimum principle.

2.9.3 Discrete Optimal Control Example

Consider the following discrete-time optimal control problem:

$$\min\ J = \sum_{t=0}^{5} \frac{1}{2}(x_t)^2 \tag{2.217}$$

subject to

$$x_{t+1} - x_t = u_t \qquad t = 0, 1, 2, 3, 4 \tag{2.218}$$

$$x_0 = 3 \tag{2.219}$$

$$-1 \leq u_t \leq 1 \quad t = 0, 1, 2, 3, 4 \tag{2.220}$$

The Hamiltonian is

$$H_t = \frac{1}{2}(x_t)^2 + \lambda_{t+1}(u_t) \quad t = 0, 1, 2, 3, 4$$

The minimum principle is

$$u_t = \begin{cases} +1 & \lambda_{t+1} < 0 \\ u_t^s & \lambda_{t+1} = 0 \\ -1 & \lambda_{t+1} > 0 \end{cases} \qquad t = 0, 1, 2, 3, 4$$

The adjoint equations are

$$\begin{aligned} \lambda_{t+1} - \lambda_t &= -\nabla_{x_t} H_t(x_t, \lambda_{t+1}, u_t) \\ &= -x_t \qquad\qquad t = 0, 1, 2, 3, 4 \end{aligned}$$

Inspection indicates that the objective function will be minimized by the application of the control $u_t = -1$ until the state variable reaches zero at an unknown time t_1; thereafter a so-called singular control $u_t^s = 0$ is applied, until the end of time horizon. Then

$$x_{t+1} - x_t = u_t = -1 \qquad t = 0, 1, \ldots, t_1$$

Since $x_0 = 3$ is given, we have

$$\begin{aligned} x_1 &= x_0 - 1 = 2 \\ x_2 &= x_1 - 1 = 1 \\ x_3 &= x_2 - 1 = 0 \end{aligned}$$

Consequently it is discovered that

$$t_1 = 2$$

Following the prior assumption, we find that

$$x_{t+1} - x_t = u_t = 0 \qquad t = 3, 4$$

which yields

$$x_4 = x_5 = 0$$

Now let us consider the conditions for adjoint variables. According to the minimum principle, we should have $\lambda_t > 0$ for $t = 0, 1, 2$ in order that $u_t = -1$ for the same time intervals. From the transversality conditions and the adjoint equations, we have

$$\begin{aligned} \lambda_5 &= 0 \\ \lambda_4 &= \lambda_5 + x_4 = 0 \\ \lambda_3 &= \lambda_4 + x_3 = 0 \\ \lambda_2 &= \lambda_3 + x_2 = 1 \\ \lambda_1 &= \lambda_2 + x_1 = 3 \\ \lambda_0 &= \lambda_1 + x_0 = 6 \end{aligned}$$

which satisfies the minimum principle. In summary, the solution is

t	0	1	2	3	4	5
x_t	+3	+2	+1	0	0	0
u_t	−1	−1	−1	0	0	0
λ_t	+6	+3	+1	0	0	0

It is instructive to approach the same problem from a purely mathematical programming perspective. In fact off-the-shelf finite-dimensional mathematical programming software or the Kuhn-Tucker conditions (without invoking the notion of the Hamiltonian and the minimum principle) may be applied directly to the nonlinear program (2.217), (2.218), (2.219), and (2.220). We leave the demonstration that the mathematical programming approach yields an identical result as an exercise for the reader.

2.10 Exercises

1. Create an example of a mathematical program with two decision variables for which no constraint qualification exists.
2. Prove or disprove: a nonlinear program with a strictly convex objective function and a non-convex feasible region arising from constraints satisfying the linear independence constraint qualification may never have a unique global optimum.
3. Solve the following nonconvex, nonlinear program graphically:

$$\min f\,(x_1, x_2) = -x_1 + 0x_2$$

subject to

$$g_1\,(x_1, x_2) = (x_1)^2 + (x_2)^2 - 2 \leq 0$$
$$g_2\,(x_1, x_2) = x_1 - (x_2)^2 \leq 0$$

4. Solve the nonconvex, nonlinear program of Exercise 3 above using the Kuhn-Tucker conditions without appeal to graphical information.
5. The example of Section 2.9.3 suggests that a *singular control* arises when it appears linearly in the Hamiltonian and has a coefficient that vanishes. Propose an alternative definition that relies on the language and optimality conditions of nonlinear programming.
6. Use the minimum principle to solve the following discrete-time optimal control problem:

$$\min J = \sum_{t=0}^{5} \left[\frac{1}{2}(x_t)^2 + u_t \right]$$

subject to

$$\begin{aligned} x_{t+1} - x_t &= u_t \qquad t = 0, 1, 2, 3, 4 \\ x_0 &= 3 \\ -1 \leq u_t &\leq 1 \quad t = 0, 1, 2, 3, 4 \end{aligned}$$

7. Use the minimum principle to solve the following discrete-time optimal control problem:

$$\min J = \sum_{t=0}^{5} \left[\frac{1}{2}(x_t)^2 + \frac{1}{2}(u_t)^2 \right]$$

subject to

$$\begin{aligned} x_{t+1} - x_t &= u_t \qquad t = 0, 1, 2, 3, 4 \\ x_0 &= 3 \\ -1 \leq u_t &\leq 1 \quad t = 0, 1, 2, 3, 4 \end{aligned}$$

List of References Cited and Additional Reading

Bazaraa, M., H. Sherali, and C. Shetty (1993). *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley.

Canon, M., C. Cullum, and E. Polak (1970). *Theory of Optimal Control and Mathematical Programming*. New York: McGraw-Hill.

Luenberger, D. G. (1984). *Linear and Nonlinear Programming*. Reading, MA: Addison-Wesley.

Mangasarian, O. (1969). *Nonlinear Programming*. New York: McGraw-Hill.

Chapter 3
Foundations of the Calculus of Variations and Optimal Control

In this chapter, we treat time as a continuum and derive optimality conditions for the extremization of certain functionals. We consider both variational calculus problems that are not expressed as optimal control problems and optimal control problems themselves. In this chapter, we relie on the classical notion of the *variation of a functional*. This classical perspective is the fastest way to obtain useful results that allow simple example problems to be solved that bolster one's understanding of continuous-time dynamic optimization.

Later in this book we will employ the more modern perspective of infinite-dimensional mathematical programming to derive the same results. The infinite-dimensional mathematical programming perspective will bring with it the benefit of shedding light on how nonlinear programming algorithms for finite-dimensional problems may be generalized and effectively applied to function spaces. In this chapter, however, we will employ the notion of a variation to derive optimality conditions in a fashion very similar to that employed by the variational-calculus pioneers.

The following is a list of the principal topics covered in this chapter:

Section 3.1: The Calculus of Variations. A formal definition of a variation is provided, along with a statement of a typical fixed-endpoint calculus of variations problem. The necessary conditions known as the Euler-Lagrange equations are derived. Other optimality conditions are also presented.

Section 3.2: Calculus of Variations Examples. Illustrative applications of the optimality conditions derived in Section 3.1 are presented.

Section 3.3: Continuous-Time Optimal Control. A cannonical optimal control problem is presented. The notion of a variation is employed to derive necessary conditions for that problem, including the Pontryagin minimum principle. Sufficiency is also discussed.

Section 3.4: Optimal Control Examples. Illustrative applications of the optimality conditions derived in Section 3.3 are presented.

Section 3.5: The Linear-Quadratic Optimal Control Problem. The linear-quadratic optimal control problem is presented; its optimality conditions are derived and employed to solve an example problem.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_3, © Springer Science+Business Media, LLC 2010

3.1 The Calculus of Varations

In this section, we take a historical approach in the spirit of the earliest investigations of the calculus of variations, assume all operations performed are valid, and avoid the formal proof of most propositions. As already indicated, our main interest is in necessary and sufficient conditions for optimality.

3.1.1 The Space $C^1\left[t_0, t_f\right]$

The space of once continuously differentiable scalar functions relative to the real interval $\left[t_0, t_f\right] \subset \Re^1_+$ is denoted by $\mathcal{C}^1\left[t_0, t_f\right]$, and we write $x \in \mathcal{C}^1\left[t_0, t_f\right]$ to denote a member of this differentiability class. When x is a vector with n components, we write $x \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n$ and say x belongs to the n-fold product of the space of once continuously differentiable functions. Although $\mathcal{C}^1\left[t_0, t_f\right]$ seems to be a sensible choice of function space, it turns out that the space of continuously differentiable functions has an important shotcoming: it is *not a complete space*. In particular, mappings defined on $\mathcal{C}^1\left[t_0, t_f\right]$ may be fail to be closed. Thus, it is possible that elementary operations involving functions belonging to $\mathcal{C}^1\left[t_0, t_f\right]$ may lead to a result that does not belong to $\mathcal{C}^1\left[t_0, t_f\right]$. However, during the early development of the calculus of variations, an alternative fundamental space for conceptualizing dynamic optimization problems was not known, and the failure of $\mathcal{C}^1\left[t_0, t_f\right]$ to be a complete space was either worked around or ignored. In our discussions, we will also encounter the space of continuous functions $\mathcal{C}^0\left[t_0, t_f\right]$, usually as the range of mappings whose domain is $\mathcal{C}^1\left[t_0, t_f\right]$. Clearly any function which belongs to $\mathcal{C}^1\left[t_0, t_f\right]$ also belongs to $\mathcal{C}^0\left[t_0, t_f\right]$.

3.1.2 The Concept of a Variation

Let us consider the abstract calculus of variations problem

$$\min J\left(x\right) = \int_{t_0}^{t_f} f_0\left[t, x\left(t\right), \frac{dx(t)}{dt}\right] dt \tag{3.1}$$

$$t_0 \text{ fixed}, x\left(t_0\right) = x_0 \text{ fixed} \tag{3.2}$$

$$t_f \text{ fixed}, x\left(t_f\right) = x^f \text{ fixed} \tag{3.3}$$

for which

$$x \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n \tag{3.4}$$

$$\frac{dx}{dt} \in \left(\mathcal{C}^0\left[t_0, t_f\right]\right)^n \tag{3.5}$$

while

$$f_0 : \Re^1_+ \times \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n \times \left(\mathcal{C}^0\left[t_0, t_f\right]\right)^n \longrightarrow \Re^1$$

Of course the initial time t_0 and the terminal time t_f are such that $t_f > t_0$ while $\left[t_0, t_f\right] \subset \Re^1_+$. If a particular curve $x(t)$ satisfies the initial conditions (3.2) as well as the terminal conditions (3.3), we say it is an *admissible trajectory*. A trajectory that maximizes or minimizes the criterion in (3.1) is called an *extremal* of $J(x)$. An admissible trajectory that minimizes $J(x)$ is a solution of the variational problem (3.1), (3.2), and (3.3). Also, the reader should note that in (3.1) the objective $J\,(x)$ should be referred to as the *criterion functional*, never as the "criterion function". This is because $J\,(x)$ is actually an operator, and we are seeking as a solution the function $x\,(t)$; this distinction is sometimes explained by saying a "functional is a function of a function."

The variation of the decision function $x\,(t)$, written as $\delta x\,(t)$, obeys

$$dx\,(t) = \delta x\,(t) + \dot{x}\,(t)\,dt \tag{3.6}$$

In other words the total differential of $x\,(t)$ is its variation $\delta x\,(t)$ *plus* the change in the variable attributed solely to time, namely $\dot{x}\,(t)\,dt$. To understand the variation of the criterion functional $J\,(x)$, we denote the change in the functional arising from the increment $h \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n$ by

$$\Delta J\,(h) \equiv J\,(x + h) - J\,(x) \tag{3.7}$$

for each $x \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n$. This allows us to make the following definition:

Definition 3.1. *Differentiability and variation of a functional. If*

$$\Delta J\,(h) = \delta J\,(h) + \varepsilon\,\|h\| \tag{3.8}$$

where, for any given $x \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n$, $\delta J\,(h)$ *is a linear functional of* $h \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n$ *and* $\varepsilon \longrightarrow 0$ *as* $\|h\| \longrightarrow 0$, $J\,(x)$ *is said to be differentiable and* $\delta J\,(h)$ *is called its variation (for the increment* h*).*

This definition is conveniently summarized by saying that the variation of the functional $J\,(x)$ is the *principal linear part of the change* $\Delta J\,(h)$. Note that the variation is dependent on the increment taken for each x. Furthermore, since the variation is the principal linear part, it may be found by retaining the linear terms of a Taylor series expansion of the criterion functional about the point x.

To illustrate let us consider the functional

$$J\,(w_1, w_2) = \int_{t_0}^{t_f} F\,(w_1, w_2)\,dt \tag{3.9}$$

$$w_1(t_0), w_2(t_0) \text{ fixed} \tag{3.10}$$

$$w_1(t_f), w_2(t_f) \text{ fixed} \tag{3.11}$$

where for convenience we take w_1, w_2 and $F(\cdot,\cdot)$ to be scalars. The change in this functional for the increment $h = (h_1, h_2)^T$ is

$$\Delta J(h_1, h_2) = J(w_1 + h_1, w_2 + h_2) - J(w_1, w_2) \tag{3.12}$$

$$= \int_{t_0}^{t_f} \left\{ F(w_1, w_2) + \frac{\partial F(w_1, w_2)}{\partial w_1} [(w_1 + h_1) - w_1] \right. \tag{3.13}$$

$$\left. + \frac{\partial F(w_1, w_2)}{\partial w_2} [(w_2 + h_2) - w_2] - F(w_1, w_2) \right\} dt + \varepsilon \|h\| \tag{3.14}$$

$$= \int_{t_0}^{t_f} \left[\frac{\partial F(w_1, w_2)}{\partial w_1} h_1 + \frac{\partial F(w_1, w_2)}{\partial w_2} h_2 \right] dt + \varepsilon \|h\| \tag{3.15}$$

It is immediate that

$$\delta J(h_1, h_2) = \int_{t_0}^{t_f} \left[\frac{\partial F(w_1, w_2)}{\partial w_1} h_1 + \frac{\partial F(w_1, w_2)}{\partial w_2} h_2 \right] dt \tag{3.16}$$

If we identify the decision variable variations δw_1 and δw_2 with the increments h_1 and h_2, respectively, this last expression becomes

$$\delta J(h_1, h_2) = \int_{t_0}^{t_f} \left[\frac{\partial F(w_1, w_2)}{\partial w_1} \delta w_1 + \frac{\partial F(w_1, w_2)}{\partial w_2} \delta w_2 \right] dt \tag{3.17}$$

which is a chain rule for the calculus of variations. Expression (3.17) is a specific instance of the following variational calculus general chain rule: the variation of the functional (3.1) obeys

$$\delta J(x) = \sum_{i=1}^{n} \int_{t_0}^{t_f} \left[\frac{\partial f_0}{\partial x_i} \delta x_i + \frac{\partial f_0}{\partial \dot{x}_i} \delta \dot{x}_i \right] dt \tag{3.18}$$

where $x(t) \in \Re^n$ for each instant of time $t \in [t_0, t_f]$. We reiterate that, in the language we have introduced, $\delta J(x)$ is the *variation of the functional* $J(x)$.

3.1.3 Fundamental Lemma of the Calculus of Variations

The necessary conditions for the classical calculus of variations depend on a specific result that herein we choose to call the *fundamental lemma of the calculus of variations*. In this section we derive that result.

In order to establish the fundamental lemma, we first state and prove a preliminary result concerning the implication of the vanishing of a certain integral. That preliminary result is the following:

Lemma 3.1. *Vanishing integral property. If* $\psi \in \mathcal{C}^0[a,b]$ *and if, for all* $\phi \in \mathcal{C}^1[a,b]$ *such that* $\phi(a) = \phi(b) = 0$*, we have*

$$\int_a^b \psi(t) \frac{d\phi}{dt}(t)\, dt = 0, \tag{3.19}$$

then $\psi(t) = c$*, a constant, for all* $t \in [a,b] \in \Re^1$.

Proof. Suppose we set

$$\phi(t) = \int_a^t [\psi(t) - c]\, dt$$

where c is defined by the relationship

$$\phi(b) = \int_a^b [\psi(t) - c]\, dt = 0$$

Note that

$$\frac{d\phi}{dt} = \psi(t) - c \tag{3.20}$$

This observation together with (3.19) tells us that

$$\begin{aligned}
0 &= \int_a^b \psi(t) \frac{d\phi}{dt}(t)\, dt = \int_a^b \psi(t)[\psi(t) - c]\, dt \\
&= \int_a^b \left\{[\psi(t)]^2 - c\psi(t)\right\} dt \\
&= \int_a^b \left\{[\psi(t)]^2 - 2c\psi(t) + c^2 + c\psi(t) - c^2\right\} dt \\
&= \int_a^b [\psi(t) - c]^2\, dt + \int_a^b c[\psi(t) - c]\, dt \\
&= \int_a^b [\psi(t) - c]^2\, dt + \int_a^b c \frac{d\phi}{dt} dt
\end{aligned}$$

Thus

$$\begin{aligned}
0 &= \int_a^b [\psi(t) - c]^2\, dt + c\int_a^b d\phi \\
&= \int_a^b [\psi(t) - c]^2\, dt + c[\phi(b) - \phi(a)] = \int_a^b [\psi(t) - c]^2\, dt
\end{aligned}$$

which can only hold if

$$\psi(t) = c \quad \forall t \in [a, b]$$

The proof is complete. ■

Now we turn to the main lemma:

Lemma 3.2. *The fundamental lemma. If* $g \in \mathcal{C}^0[a, b]$, $h \in \mathcal{C}^0[a, b]$ *and if, for all* $\phi \in \mathcal{C}^1[a, b]$ *such that* $\phi(a) = \phi(b) = 0$, *we have*

$$\int_a^b \left[g(t)\phi(t) + h(t)\dot{\phi}(t)\right] dt = 0,$$

then

$$g(t) = \frac{dh(t)}{dt} \quad \forall t \in [a, b]$$

Proof. Define

$$G(t) = \int_a^t g(\tau)\, d\tau$$

and consider the integral

$$\int_a^b G(t)\, d\phi(t) = \int_a^b G(t) \frac{d\phi}{dt} dt \tag{3.21}$$

Using the standard formula for integration by parts, (3.21) can be stated as

$$\int_a^b G(t) \frac{d\phi}{dt} dt = [G(t)\phi(t)]_a^b - \int_a^b g(t)\phi(t)\, dt \tag{3.22}$$

where this last result holds for all $\phi \in \mathcal{C}^1[a, b]$ such that $\phi(a) = \phi(b) = 0$ per the given. It is immediate that

$$[G(t)\phi(t)]_a^b = G(b)\phi(b) - G(a)\phi(a) = 0$$

so that (3.22) becomes

$$\int_a^b G(t) \frac{d\phi}{dt} dt = -\int_a^b g(t)\phi(t)\, dt \tag{3.23}$$

for all $\phi \in \mathcal{C}^1[a, b]$. Consequently, we have

$$\int_a^b \left[g(t)\phi(t) + h(t)\dot{\phi}(t)\right] dt = \int_a^b \left[-G(t) + h(t)\right]\dot{\phi}(t)\, dt = 0 \tag{3.24}$$

By the Lemma 3.1, it follows that

$$-G(t) + h(t) = c, \text{ a constant } \forall t \in [a, b]$$

Hence

$$g(t) = \frac{dG(t)}{dt} = \frac{dh(t)}{dt} \tag{3.25}$$

which complete the proof. ■

We note that Lemma 3.2 is easily generalized to deal with $\phi \in \left(\mathcal{C}^1[a, b]\right)^n$.

3.1.4 Derivation of the Euler-Lagrange Equation

In this section, we derive necessary conditions for the following calculus of variations problem:

$$\min J(x) = \int_{t_0}^{t_f} f_0\left[t, x(t), \frac{dx}{dt}(t)\right] dt \tag{3.26}$$

$$t_0 \text{ fixed}, x(t_0) = x_0 \text{ fixed} \tag{3.27}$$

$$t_f \text{ fixed}, x\left(t_f\right) = x^f \text{ fixed} \tag{3.28}$$

where $x \in \left(\mathcal{C}^1\left[t_0, t_f\right]\right)^n$ is the decision function we are seeking. In what follows, we interpret partial derivative operators to be gradients when $n \geq 2$; that is

$$\frac{\partial f_0(t, x, \dot{x})}{\partial x} = \left[\nabla_x f_0(t, x, \dot{x})\right]^T$$

$$\frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}} = \left[\nabla_{\dot{x}} f_0(t, x, \dot{x})\right]^T$$

Next we construct the variation $\delta J(x)$ by taking the principal linear part of

$$\Delta J = J(x + h) - J(x) \tag{3.29}$$

$$= \int_{t_0}^{t_f} f_0\left(t, x + h, \frac{d(x + h)}{dt}\right) dt - \int_{t_0}^{t_f} f_0\left(t, x, \frac{dx}{dt}\right) dt \tag{3.30}$$

$$= \int_{t_0}^{t_f} \left[f_0\left(t, x + h, \frac{d(x + h)}{dt}\right) - f_0\left(t, x, \frac{dx}{dt}\right)\right] dt \tag{3.31}$$

where h is an arbitrary increment. Making a Taylor series expansion of the integrand about the point $(x, \dot{x})$, this becomes

$$\Delta J = \int_{t_0}^{t_f} \left[f_0\left(t, x, \frac{dx}{dt}\right) - f_0\left(t, x, \frac{dx}{dt}\right) \right] dt$$
$$+ \int_{t_0}^{t_f} \left\{ \frac{\partial f_0\,(t, x, \dot{x})}{\partial x} [(x + h) - x] \right.$$
$$\left. + \frac{\partial f_0\,(t, x, \dot{x})}{\partial \dot{x}} \left[\left(\dot{x} + \dot{h}\right) - \dot{x} \right] \right\} dt + \ldots$$

which tells us that

$$\delta J\,(x) = \int_{t_0}^{t_f} \left[\frac{\partial f_0\,(t, x, \dot{x})}{\partial x} h + \frac{\partial f_0\,(t, x, \dot{x})}{\partial \dot{x}} \dot{h} \right] dt \tag{3.32}$$

A necessary condition for our original problem is

$$\delta J\,(x) = 0$$

Thus, upon invoking the fundamental lemma of the calculus of variations (Lemma 3.2), we see that

$$\frac{\partial f_0\,(t, x, \dot{x})}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0\,(t, x, \dot{x})}{\partial \dot{x}} \right] = 0 \tag{3.33}$$
$$x\,(t_0) = x_0 \tag{3.34}$$
$$x\left(t_f\right) = x^f \tag{3.35}$$

where (3.33) is called the Euler-Lagrange equation, and (3.34) and (3.35) are the original boundary conditions of the problem we have analyzed; note in particular that these relationships constitute necessary conditions for the calculus of variations problem (3.26), (3.27), and (3.28).

We turn now to another form of the Euler-Lagrange equation. Note that an application of the chain rule yields the following two results:

$$\frac{df_0}{dt} = \frac{\partial f_0}{\partial t} + \frac{\partial f_0}{\partial x}\frac{dx}{dt} + \frac{\partial f_0}{\partial \dot{x}}\frac{d\dot{x}}{dt}$$
$$\frac{d}{dt}\left(\dot{x}\frac{\partial f_0}{\partial \dot{x}}\right) = \frac{\partial f_0}{\partial \dot{x}}\frac{d\dot{x}}{dt} + \dot{x}\frac{d}{dt}\left(\frac{\partial f_0}{\partial \dot{x}}\right)$$

Combining the above two equations, we get

$$\frac{d}{dt}\left(\dot{x}\frac{\partial f_0}{\partial \dot{x}}\right) = \left[\frac{df_0}{dt} - \frac{\partial f_0}{\partial t} - \frac{\partial f_0}{\partial x}\frac{dx}{dt}\right] + \dot{x}\frac{d}{dt}\left(\frac{\partial f_0}{\partial \dot{x}}\right)$$

Upon reordering

$$\dot{x}\left[\frac{\partial f_0}{\partial x}-\frac{d}{dt}\left(\frac{\partial f_0}{\partial \dot{x}}\right)\right]+\frac{d}{dt}\left(\dot{x}\frac{\partial f_0}{\partial \dot{x}}\right)-\frac{df_0}{dt}+\frac{\partial f_0}{\partial t}=0 \tag{3.36}$$

Note that the first term of (3.36) contains the righthand side of the Euler-Lagrange equation. Therefore, if x is a solution of the Euler-Lagrange equation, then

$$\frac{d}{dt}\left(\dot{x}\frac{\partial f_0}{\partial \dot{x}}-f_0\right)+\frac{\partial f_0}{\partial t}=0 \tag{3.37}$$

Result (3.37), known as *the second form of the Euler-Lagrange equation*, can be very useful in particular cases. Note that if $\dfrac{\partial f_0}{\partial t}=0$, then

$$\frac{d}{dt}\left(\dot{x}\frac{\partial f_0}{\partial \dot{x}}-f_0\right)=0 \quad \text{or} \quad \dot{x}\frac{\partial f_0}{\partial \dot{x}}-f_0=\text{constant}$$

We will study an application of the second form later in this chapter.

3.1.5 Additional Necessary Conditions in the Calculus of Variations

In certain situations the first-order necessary conditions are not adequate to fully describe an extremal trajectory, and additional necessary conditions are needed. In particular, it is possible to derive second-order necessary conditions. To that end, we need to first establish the following result:

Lemma 3.3. *Nonnegativity of a functional. Let* $g\in\mathcal{C}^0[a,b]$, $h\in\mathcal{C}^0[a,b]$, *and* $\phi\in\mathcal{C}^1[a,b]$; *suppose also that* $\phi(a)=\phi(b)=0$. *A necessary condition for the functional*

$$F=\int_a^b\left[g(t)\{\phi(t)\}^2+h(t)\left\{\dot{\phi}(t)\right\}^2\right]dt \tag{3.38}$$

to be nonnegative for all ϕ *is that*

$$h(t)\geq 0 \quad \forall t\in[a,b]$$

Proof. See Gelfand and Fomin (2000). ■

For a function $f(x)$ to be minimized, we know from Chapter 2 that a second-order necessary condition is $f''(x)\geq 0$ at the minimum. For the problem (3.26),

(3.27), and (3.28), we have a similar variational calculus necessary condition called *Legendre's condition*, which is given in the following theorem:

Theorem 3.1. *Legendre's condition. A necessary condition for x to minimize $J(x)$ in the problem defined by (3.26),(3.27), and (3.28) is that*

$$\frac{\partial^2 f_0\left(t, x, \dot{x}\right)}{\partial \dot{x}^2} \geq 0$$

for all $t \in \left[t_0, t_f\right]$. When maximizing, the inequality is reversed.

Proof. Let us define

$$I(\epsilon) = J(x + \epsilon\phi) = \int_{t_0}^{t_f} f_0\left[t, x + \epsilon\phi, \dot{x} + \epsilon\dot{\phi}\right] dt$$

for all $\phi \in \mathcal{C}^1\left[t_0, t_f\right]$ such that $\phi\left(t_0\right) = \phi\left(t_f\right) = 0$. For $J(x)$ to be minimized at x, $I(\epsilon)$ should be minimized at $\epsilon = 0$. That is

$$\frac{d^2}{d\epsilon^2} I(\epsilon)\bigg|_{\epsilon=0} \geq 0$$

Note that

$$\begin{aligned}
\frac{d}{d\epsilon} I(\epsilon)\bigg|_{\epsilon=0} &= \int_{t_0}^{t_f} \left[\frac{\partial f_0}{\partial x}\phi + \frac{\partial f_0}{\partial \dot{x}}\dot{\phi}\right] dt \\
\frac{d^2}{d\epsilon^2} I(\epsilon)\bigg|_{\epsilon=0} &= \int_{t_0}^{t_f} \left[\frac{\partial^2 f_0}{\partial x^2}\phi^2 + \frac{\partial^2 f_0}{\partial x \partial \dot{x}}\phi\dot{\phi} + \frac{\partial^2 f_0}{\partial \dot{x}^2}\dot{\phi}^2 + \frac{\partial^2 f_0}{\partial x \partial \dot{x}}\phi\dot{\phi}\right] dt \\
&= \int_{t_0}^{t_f} \left[\frac{\partial^2 f_0}{\partial x^2}\phi^2 + 2\frac{\partial^2 f_0}{\partial x \partial \dot{x}}\phi\dot{\phi} + \frac{\partial^2 f_0}{\partial \dot{x}^2}\dot{\phi}^2\right] dt \qquad (3.39)
\end{aligned}$$

Integrating by parts, we have

$$\begin{aligned}
\int_{t_0}^{t_f} 2\frac{\partial^2 f_0}{\partial x \partial \dot{x}}\phi\dot{\phi}\, dt &= \frac{\partial^2 f_0}{\partial x \partial \dot{x}}\phi^2\bigg|_{t_0}^{t_f} - \int_{t_0}^{t_f} \frac{d}{dt}\left[\frac{\partial^2 f_0}{\partial x \partial \dot{x}}\right]\phi^2 dt \\
&= -\int_{t_0}^{t_f} \frac{d}{dt}\left[\frac{\partial^2 f_0}{\partial x \partial \dot{x}}\right]\phi^2 dt \qquad (3.40)
\end{aligned}$$

where we have used the boundary conditions $\phi\left(t_0\right) = \phi\left(t_f\right) = 0$. Substituting (3.40) into (3.39), we get

$$\frac{d^2}{d\epsilon^2} I(\epsilon)\bigg|_{\epsilon=0} = \int_{t_0}^{t_f} \left[\left(\frac{\partial^2 f_0}{\partial x^2} - \frac{d}{dt}\left[\frac{\partial^2 f_0}{\partial x \partial \dot{x}}\right]\right)\phi^2 + \frac{\partial^2 f_0}{\partial \dot{x}^2}\dot{\phi}^2\right] dt \qquad (3.41)$$

For (3.41) to be nonnegative, we must have, by Lemma 3.3

$$\frac{\partial^2 f_0}{\partial \dot{x}^2} \geq 0$$

for all $t \in \left[t_0, t_f\right]$. This completes the proof. ■

We now turn our attention to another necessary condition, namely the so-called *Weierstrass condition*, which is the subject of the following theorem:

Theorem 3.2. *Weierstrass condition. For the problem defined by (3.26), (3.27), and (3.28), if $x(t)$ is the solution, then we have*

$$E(t, x, \dot{x}, \dot{y}) \geq 0$$

where $E(\cdot)$ is the Weierstrass excess function (E-function)

$$E(t, x, \dot{x}, \dot{y}) = f_0(t, x, \dot{y}) - f_0(t, x, \dot{x}) - \frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}}(\dot{y} - \dot{x}) \tag{3.42}$$

for all admissible y.

Proof. We follow Spruck (2006). Let us pick $t_1 \in [t_0, t_f]$ and $\dot{y}$ to be fixed but otherwise arbitrary. For both ϵ and h positive, fixed and suitably small, we define

$$\eta(t) = \begin{cases} t - t_1 & t \in [t_1, t_1 + \epsilon h] \\ \dfrac{\epsilon}{1-\epsilon}(t_1 + h - t) & t \in [t_1 + \epsilon h, t_1 + h] \\ 0 & \text{otherwise} \end{cases}$$

We also define

$$x^1 = x + \eta(y - \dot{x}(t_1)) \tag{3.43}$$

Since x is a minimizer

$$J(x) \leq J(x^1)$$

As a consequence we have

$$0 \leq J(x^1) - J(x)$$

It then follows that

$$0 \leq \int_{t_1}^{t_1+h} f_0\left(t, x(t) + \eta(\dot{y} - \dot{x}(t_1)), \dot{x}(t) + \dot{\eta}(\dot{y} - \dot{x}(t_1)\right) dt - \int_{t_1}^{t_1+h} f_0(t, x(t), \dot{x}(t)) dt$$

$$= \int_{t_1}^{t_1+\epsilon h} f_0(t, x(t) + (t-t_1)(\dot{y} - \dot{x}(t_1)), \dot{x} + \dot{y} - \dot{x}(t_1))dt$$
$$+ \int_{t_1+\epsilon h}^{t_1+h} f_0(t, x(t) + \frac{\epsilon}{1-\epsilon}(t_1 + h - t)(\dot{y} - \dot{x}(t_1)), \dot{x} - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1)))dt$$
$$- \int_{t_1}^{t_1+h} f_0(t, x, \dot{x})dt \tag{3.44}$$

We wish to simplify (3.44); this is accomplished by introducing the change of variable

$$t = t_1 + \xi h$$

which leads directly to

$$0 \leq A + B + C \tag{3.45}$$

where

$$A = h \int_0^{\epsilon} f_0(t_1 + \xi h, x(t_1 + \xi h) + \xi h(\dot{y} - \dot{x}(t_1)), \dot{x}(t_1 + \xi h) + \dot{y} - \dot{x}(t_1))d\xi \tag{3.46}$$

$$B = h \int_{\epsilon}^{1} f_0(t_1 + \xi h, x(t_1 + \xi h) + \frac{\epsilon}{1-\epsilon}(1-\xi)h(\dot{y} - \dot{x}(t_1)), \dot{x}(t_1 + \xi h)$$
$$- \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1)))d\xi \tag{3.47}$$

$$C = -h \int_0^{1} f_0(t_1 + \xi h, x(t_1 + \xi h), \dot{x}(t_1 + \xi h))d\xi \tag{3.48}$$

Dividing by h and taking the limit $h \to 0$, for each of the terms above, we obtain

$$0 \leq \int_0^{\epsilon} f_0(t_1, x(t_1), \dot{y})d\xi + \int_{\epsilon}^{1} f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1)))d\xi$$
$$- \int_0^{1} f_0(t_1, x(t_1), \dot{x}(t_1))d\xi \tag{3.49}$$

Because the integrands of (3.49) are independent of ξ we may write

$$0 \leq \epsilon f_0(t_1, x(t_1), \dot{y}) + (1-\epsilon) f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1)))$$
$$- f_0(t_1, x(t_1), \dot{x}(t_1)) \tag{3.50}$$

Dividing (3.50) by ϵ yields

$$0 \leq f_0(t_1, x(t_1), \dot{y}) + \frac{1-\epsilon}{\epsilon} f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1)))$$
$$- \frac{1}{\epsilon} f_0(t_1, x(t_1), \dot{x}(t_1))$$

$$
\begin{aligned}
&= f_0(t_1, x(t_1), \dot{y}) - f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1))) \\
&\quad + \frac{1}{\epsilon}\left[f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1))) - f_0(t_1, x(t_1), \dot{x}(t_1)) \right] \\
&= f_0(t_1, x(t_1), \dot{y}) - f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1))) \\
&\quad - \frac{f_0(t_1, x(t_1), \dot{x}(t_1)) - f_0(t_1, x(t_1), \dot{x}(t_1) - \frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1)))}{\frac{\epsilon}{1-\epsilon}(\dot{y} - \dot{x}(t_1))} \cdot \frac{\dot{y} - \dot{x}(t_1)}{1-\epsilon}
\end{aligned}
\tag{3.51}
$$

Taking the limit of (3.51) as $\epsilon \to 0$, we get

$$
0 \leq f_0(t_1, x(t_1), \dot{y}) - f_0(t_1, x(t_1), \dot{x}(t_1)) - \frac{\partial f_0(t_1, x(t_1), \dot{x}(t_1))}{\dot{x}(t_1)}(\dot{y} - \dot{x}(t_1))
$$

Since $\dot{y}$ and t_1 are arbitrary, the theorem is proven. ■

Reflection on the apparatus introduced above reveals that we may consider any solution trajectory to be a continuous function of time that is *piecewise smooth.* We note, however, that when an admissible function $x(t)$ is piecewise smooth, $\dot{x}(t)$ need not be continuous but rather only piecewise continuous. This is because a piecewise smooth curve may have points, often loosely referred to as *corners*, at which the first derivative is discontinuous. Specifically, it exhibits a jump discontinuity. For such points of jump discontinuity of the time deivative (corners), we have necessary conditions, called *Weierstrass-Erdman conditions*, which are the subject of the following result:

Theorem 3.3. *Weierstrass-Erdman conditions. For the problem defined by (3.26),(3.27), and (3.28), suppose an optimal solution $x(t)$ has a jump discontinuity of its time derivative at $t = t_1$. Then the Weierstrass-Erdman conditions*

$$
\left[\frac{\partial f_0}{\partial \dot{x}}\right]_{t=t_1^-} = \left[\frac{\partial f_0}{\partial \dot{x}}\right]_{t=t_1^+}
$$

$$
\left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^-} = \left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^+}
$$

must hold.

Proof. We follow Gelfand and Fomin (2000) and observe that for $J(x)$ to have a minimum, the Euler equation must be satisfied:

$$
\frac{\partial f_0}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0}{\partial \dot{x}}\right] = 0 \tag{3.52}
$$

Let us decompose the objective functional so that $J = J_1 + J_2$ where

$$J_1 = \int_{t_0}^{t_1} f_0(t, x, \dot{x})\, dt$$
$$J_2 = \int_{t_1}^{t_f} f_0(t, x, \dot{x})\, dt$$

and $t_1 \in [t_0, t_f]$. Since $x(t_1)$ is not a fixed point, we have

$$\begin{aligned}\Delta J_1 &= J_1(x+h) - J_1(x) \\ &= \int_{t_0}^{t_1+\delta t_1} f_0\left(t, x+h, \dot{x}+\dot{h}\right) dt - \int_{t_0}^{t_1} f_0(t, x, \dot{x})\, dt\end{aligned}$$

$$= \int_{t_0}^{t_1} \left\{ f_0\left(t, x+h, \dot{x}+\dot{h}\right) - f_0(t, x, \dot{x}) \right\} dt + \int_{t_1}^{t_1+\delta t_1} f_0\left(t, x+h, \dot{x}+\dot{h}\right) dt$$

$$= \int_{t_0}^{t_1} \left\{ \frac{\partial f_0}{\partial x} h + \frac{\partial f_0}{\partial \dot{x}} \dot{h} \right\} dt + \varepsilon \|h\| + [f_0]_{t=t_1} \delta t_1 + \varepsilon \|\delta t_1\|$$

where we have employed a Taylor series expansion. It is immediate that the principal linear part of the above expansion is

$$\delta J_1 = \int_{t_0}^{t_1} \left\{ \frac{\partial f_0}{\partial x} h + \frac{\partial f_0}{\partial \dot{x}} \dot{h} \right\} dt + [f_0]_{t=t_1} \delta t_1 \tag{3.53}$$

Integrating by parts we get

$$\delta J_1 = \int_{t_0}^{t_1} \left\{ \frac{\partial f_0}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0}{\partial \dot{x}}\right] \right\} h dt + \left[\frac{\partial f_0}{\partial \dot{x}} h\right]_{t=t_1} + [f_0]_{t=t_1} \delta t_1$$

From (3.52), we have

$$\delta J_1 = \left[\frac{\partial f_0}{\partial \dot{x}} h\right]_{t=t_1^-} + [f_0]_{t=t_1^-} \delta t_1^-$$

Since h is arbitrary, we are free to set

$$h(t_1^-) = \delta x(t_1^-) - \dot{x}(t_1^-) \delta t_1^-$$

provided δx is arbitrary. Therefore

$$\delta J_1 = \left[\frac{\partial f_0}{\partial \dot{x}}\delta x\right]_{t=t_1^-} + \left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^-} \delta t_1^-$$

Similarly

$$\delta J_2 = -\left[\frac{\partial f_0}{\partial \dot{x}}\delta x\right]_{t=t_1^+} - \left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^+} \delta t_1^+$$

Continuity for $x(t)$ at $t = t_1$ implies

$$\begin{aligned} \delta x(t_1^-) &= \delta x(t_1^+) \\ \delta t_1^- &= \delta t_1^+ \end{aligned}$$

So we must have

$$\begin{aligned} 0 = \delta J &= \delta J_1 + \delta J_1 \\ &= \left(\left[\frac{\partial f_0}{\partial \dot{x}}\right]_{t=t_1^-} - \left[\frac{\partial f_0}{\partial \dot{x}}\right]_{t=t_1^+}\right)\delta x(t_1^-) \\ &\quad + \left(\left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^-} - \left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^+}\right)\delta t_1^- \end{aligned}$$

Since $\delta x(t_1^-)$ and δt_1^- are arbitrary, the conditions

$$\begin{aligned} \left[\frac{\partial f_0}{\partial \dot{x}}\right]_{t=t_1^-} &= \left[\frac{\partial f_0}{\partial \dot{x}}\right]_{t=t_1^+} \\ \left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^-} &= \left[f_0 - \frac{\partial f_0}{\partial \dot{x}}\dot{x}\right]_{t=t_1^+} \end{aligned}$$

must hold. This completes the proof. ■

3.1.6 Sufficiency in the Calculus of Variations

In this section, we derive sufficient conditions for optimality in variational calculus problems. We begin by stating and proving the following result:

Theorem 3.4. *Sufficiency of the Euler-Lagrange equation. Consider the variational calculus problem defined by (3.26),(3.27), and (3.28). Let the integrand f_0 be convex with respect to $(x, \dot{x})$ for each instant of time t considered. Furthermore, let $x^*(t)$ be a piecewise smooth, admissible function satisfying the Euler-Lagrange equation*

everywhere except possibly at points of jump discontinuity of its time derivative where it satisfies the Weierstrass-Erdman conditions. Then $x^*(t)$ is a solution to (3.26),(3.27), and (3.28).

Proof. We follow Brechtken-Manderscheid (1991). Because of the assumed convexity, we have

$$f_0(t,x,\dot{x}) \geq f_0\left(t,x^*,\dot{x}^*\right) + \frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial x}\left(x-x^*\right) + \frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial \dot{x}}\left(\dot{x}-\dot{x}^*\right) \tag{3.54}$$

Now let x and x^* be two piecewise smooth admissible functions, and let the corners of x and x^* correspond to instants of time t_i for $i \in [1,m]$ such that

$$t_0 < t_1 < \cdots < t_m < t_{m+1} = t_f$$

Using (3.54), we may write

$$\begin{aligned} J(x) - J\left(x^*\right) &= \int_{t_0}^{t_f} \left[f_0(t,x,\dot{x}) - f_0\left(t,x^*,\dot{x}^*\right)\right] dt \\ &\geq \int_{t_0}^{t_f} \left[\frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial x}\left(x-x^*\right) + \frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial \dot{x}}\left(\dot{x}-\dot{x}^*\right)\right] dt \end{aligned} \tag{3.55}$$

Integrating by parts, we have

$$\begin{aligned} \int_{t_0}^{t_f} \frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial \dot{x}}\left(\dot{x}-\dot{x}^*\right) dt = &\sum_{i=1}^{m+1} \left[\frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial \dot{x}}\left(x-x^*\right)\right]_{t=t_{i-1}}^{t=t_i} \\ &- \int_{t_0}^{t_f} \frac{d}{dt}\left[\frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial \dot{x}}\left(x-x^*\right)\right] dt \end{aligned} \tag{3.56}$$

Since x^* satisfies the Weierstrass-Erdman corner conditions while x and x^* are identical at the both endpoints, the first term on the righthand side of (3.56) is zero. Therefore, we have

$$J(x) - J\left(x^*\right) \geq \int_{t_0}^{t_f} \left\{\frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0(t,x^*,\dot{x}^*)}{\partial \dot{x}}\right]\right\}\left(x-x^*\right) dt = 0$$

because of the assumption that x^* satisfies the Euler-Lagrange equation. This completes the proof. ■

3.1.7 Free Endpoint Conditions in the Calculus of Variations

Now consider the problem

$$\min J(x) = \int_{t_0}^{t_f} f_0\left[t, x(t), \frac{dx}{dt}(t)\right] dt \tag{3.57}$$

$$t_0 \text{ fixed}, x(t_0) \text{ free} \tag{3.58}$$

$$t_f \text{ fixed}, x\left(t_f\right) \text{ free} \tag{3.59}$$

where the endpoints are free. Clearly, the present circumstances require invocation of boundary conditions different from those used for the fixed endpoint problem. In particular, the boundary conditions are chosen to make the variation $\delta J(x)$ expressed by

$$\begin{aligned}\delta J &= \int_{t_0}^{t_f} \left\{\frac{\partial f_0}{\partial x}h + \frac{\partial f_0}{\partial \dot{x}}\dot{h}\right\} dt \\ &= \int_{t_0}^{t_f} \left\{\frac{\partial f_0}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0}{\partial \dot{x}}\right]\right\} h dt + \left[\frac{\partial f_0}{\partial \dot{x}}h\right]_{t=t_0}^{t=t_f}\end{aligned}$$

vanish. This is accomplished by enforcing the Euler-Lagrange equation plus the conditions

$$\left[\frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}}\right]_{t_f} = 0 \tag{3.60}$$

$$\left[\frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}}\right]_{t_0} = 0 \tag{3.61}$$

which are the *free endpoint conditions*; sometimes they are also called the *natural boundary conditions*. Note that they are to be enforced only when the endpoints are free. Furthermore, if only one endpoint is free then only the associated free endpoint condition is enforced.

3.1.8 Isoperimetric Problems in the Calculus of Variations

In the calculus of variations, one may encouter constraints, over and above endpoint conditions, that must be satisfied by an admissible trajectory. An important class of such constraints is a type of integral constraint that is often referred to as an

isoperimetric constraint; the presence of such a constraint causes the problem of interest to take the following form:

$$\min\ J(x) = \int_{t_0}^{t_f} f_0(t, x, \dot{x})dt \tag{3.62}$$

$$x(t_0) = x_0 \tag{3.63}$$

$$x(t_f) = x^f \tag{3.64}$$

$$K(x) = \int_{t_0}^{t_f} g(t, x, \dot{x})dt = c \tag{3.65}$$

where $g \in \mathcal{C}^1\left[t_0, t_f\right]$ and c is a constant. Necessary conditions for this problem are provided by the following theorem:

Theorem 3.5. *Isoperimetric constraints and the Euler-Lagrange equation. Let the problem defined by (3.62),(3.63),(3.64), and (3.65) have a minimum at $x(t)$. Then if $x(t)$ is not an extremal of $K(x)$, there exists a constant λ such that $x(t)$ is a minimizer of the functional*

$$\int_{t_0}^{t_f} (f_0 + \lambda g)\, dt \tag{3.66}$$

That is, $x(t)$ satisfies the Euler-Lagrange equation for (3.66):

$$\frac{\partial f_0}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0}{\partial \dot{x}}\right] + \lambda\left(\frac{\partial g}{\partial x} - \frac{d}{dt}\left[\frac{\partial g}{\partial \dot{x}}\right]\right) = 0$$

This theorem is analogous to the Lagrange multiplier rule for mathamatical programs with equality constraints; hence, its proof is not given here, but, instead, is left as an exercise for the reader. Intriligator (1971) points out a duality-type result for the isoperimetric problem, which he calls the *principle of reciprocity*. It states that if $x(t)$ minimizes J subject to the condition that K is constant, then for certain mild regularity conditions $x(t)$ maximizes K subject to the condition that J is a constant.

3.1.9 The Beltrami Identity for $\frac{\partial f_0}{\partial t} = 0$

We now derive a result that is useful for studying the Brachistochrone problem, a numerical example of which we shall shortly consider. For now consider the chain rule

$$\frac{df_0}{dt} = \frac{\partial f_0}{\partial x}\dot{x} + \frac{\partial f_0}{\partial \dot{x}}\frac{d\dot{x}}{dt} + \frac{\partial f_0}{\partial t} \tag{3.67}$$

which may be re-expressed as

$$\frac{\partial f_0}{\partial x}\dot{x} = \frac{df_0}{dt} - \frac{\partial f_0}{\partial \dot{x}}\frac{d\dot{x}}{dt} - \frac{\partial f_0}{\partial t} \tag{3.68}$$

Consider now the expression

$$\dot{x}\frac{\partial f_0}{\partial x} = \dot{x}\frac{d}{dt}\left[\frac{\partial f_0}{\partial \dot{x}}\right] \tag{3.69}$$

which is the Euler-Lagrange equation multiplied by $\dot{x}$. Substituting (3.69) into (3.68) we obtain

$$\frac{df_0}{dt} - \frac{\partial f_0}{\partial \dot{x}}\frac{d\dot{x}}{dt} - \frac{\partial f_0}{\partial t} - \dot{x}\frac{d}{dt}\left[\frac{\partial f_0}{\partial \dot{x}}\right] = 0 \tag{3.70}$$

We note that

$$\frac{d}{dt}\left[f_0 - \dot{x}\frac{\partial f_0}{\partial \dot{x}}\right] = \frac{df_0}{dt} - \frac{\partial f_0}{\partial \dot{x}}\frac{d\dot{x}}{dt} - \dot{x}\frac{d}{dt}\frac{\partial f_0}{\partial \dot{x}} \tag{3.71}$$

so

$$-\frac{\partial f_0}{\partial t} + \frac{d}{dt}\left[f_0 - \dot{x}\frac{\partial f_0}{\partial \dot{x}}\right] = 0 \tag{3.72}$$

When the variational problem of interest has an integrand f_0 that is independent of time, that is when

$$\frac{\partial f_0}{\partial t} = 0\,, \tag{3.73}$$

it is immediate that

$$\frac{d}{dt}\left[f_0 - \dot{x}\frac{\partial f_0}{\partial \dot{x}}\right] = 0 \tag{3.74}$$

In other words

$$\dot{x}\frac{\partial f_0}{\partial \dot{x}} - f_0 = C_0\,, \text{ a constant} \tag{3.75}$$

Expression (3.75) is Beltrami's identity.

3.2 Calculus of Variations Examples

Next we provide some solved examples that make use of the calculus of variations optimality conditions discussed previously.

3.2.1 Example of Fixed Endpoints in the Calculus of Variations

Consider

$$\begin{aligned} f_0(x, \dot{x}, t) &= \frac{1}{2}\left[x^2 + \left(\frac{dx}{dt}\right)^2\right] \\ t_0 &= 0 \\ t_f &= 5 \\ x(t_0) &= 10 \\ x(t_f) &= 0 \end{aligned}$$

That is, we wish to solve

$$\begin{aligned} \min J(x) &= \int_{t_0}^{t_f} \frac{1}{2}\left[x^2 + \left(\frac{dx}{dt}\right)^2\right] dt \\ x(0) &= 10 \\ x(5) &= 0 \end{aligned}$$

Note that

$$\begin{aligned} \frac{\partial f_0(t, x, \dot{x})}{\partial x} &= x \\ \frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}} &= \frac{dx}{dt} \\ \frac{d}{dt}\left[\frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}}\right] &= \frac{d^2x}{dt^2} \end{aligned}$$

Therefore

$$\frac{\partial f_0(t, x, \dot{x})}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0(t, x, \dot{x})}{\partial \dot{x}}\right] = x - \frac{d^2x}{dt^2}$$

and the Euler-Lagrange equation with endpoint conditions is

$$\begin{aligned} \frac{d^2x}{dt^2} - x &= 0 \\ x(0) &= 10 \\ x(5) &= 0 \end{aligned}$$

The exact solution is

$$x(t) = \frac{1}{e^t e^5 - e^t e^{-5}}\left(10e^5 - 10e^{-5}e^{2t}\right) \tag{3.76}$$

We leave as an exercise for the reader the determination of whether this particular problem also satisfies the conditions that assure the Euler-Lagrange equation is sufficient and allow us to determine (3.76) is in fact an optimal solution.

3.2.2 Example of Free Endpoints in the Calculus of Variations

Consider

$$\begin{aligned} f_0(t,x,\dot{x}) &= \frac{1}{2}\left[x+\left(\frac{dx}{dt}\right)^2\right] \\ t_0 &= 0 \\ t_f &= 5 \\ x(t_0) &= 5 \\ x(t_f) &\text{ free} \end{aligned}$$

That is, we wish to solve

$$\begin{aligned} \min J(x) &= \int_0^5 \frac{1}{2}\left[x+\left(\frac{dx}{dt}\right)^2\right] dt \\ x(0) &= 5 \\ x(5) &\text{ free} \end{aligned}$$

Note that

$$\begin{aligned} \frac{\partial f_0(t,x,\dot{x})}{\partial x} &= \frac{1}{2} \\ \frac{\partial f_0(t,x,\dot{x})}{\partial \dot{x}} &= \frac{dx}{dt} \\ \frac{d}{dt}\left[\frac{\partial f_0(t,x,\dot{x})}{\partial \dot{x}}\right] &= \frac{d^2x}{dt^2} \end{aligned}$$

and, therefore, the Euler-Lagrange equation is

$$\frac{\partial f_0(t,x,\dot{x})}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0(t,x,\dot{x})}{\partial \dot{x}}\right] = \frac{1}{2} - \frac{d^2x}{dt^2} = 0$$

Since we have a free endpoint, the terminal condition is

$$\left[\frac{\partial f_0(t,x,\dot{x})}{\partial \dot{x}}\right]_{t=t_f} = \left[\frac{dx}{dt}\right]_{t=5} = 0$$

Thus, we have the following ordinary differential equation with boundary conditions:

$$\frac{d^2x}{dt^2} = \frac{1}{2} \tag{3.77}$$

$$x(0) = 5 \tag{3.78}$$

$$\frac{dx(5)}{dt} = 0 \tag{3.79}$$

As may be verified by direct substitution, the exact solution is

$$x(t) = \frac{1}{4}t^2 - \frac{5}{2}t + 5 \tag{3.80}$$

3.2.3 The Brachistochrone Problem

In Chapter 1 we encountered the famous *brachistochrone problem*, which is generally thought to have given birth to the branch of mathematics that is known as the calculus of variations. Recall that we used $y\,(x)$ to denote the vertical position of a bead sliding on a wire as a function of its horizontal position x. What we now want to do is extremize the functional $J\,(x, y, y')$ associated with that problem. To that end, we consider the start point $P_A = (x_A, y_A) = (0, 0)$ and the endpoint $P_B = (x_B, y_B)$. Let us define a variable s to denote arc length along the wire; then, a segment of the wire must obey

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + (y')^2}dx$$

where

$$y' = \frac{dy}{dx} \tag{3.81}$$

Assuming a constant gravitaional acceleration g and invoking conservation of energy, we see the speed of the bead, v, obeys

$$\frac{1}{2}mv^2 - mgy = 0$$

which in turn requires

$$v = \sqrt{2gy} \tag{3.82}$$

The total travel time J, which is to be minimized, may be stated as

$$\begin{aligned} J &= \int_{P_A}^{P_B} \frac{ds}{v} \\ &= \int_0^{x_B} \frac{\sqrt{1 + (y')^2}dx}{\sqrt{2gy}} \\ &= \frac{1}{\sqrt{2g}} \int_0^{x_B} \frac{\sqrt{1 + (y')^2}}{\sqrt{y}}dx \end{aligned}$$

Therefore, our problem may be given the form

$$\min J\left(x, y, y'\right) = \frac{1}{(2g)^{\frac{1}{2}}} \int_0^{x_B} \left[1 + \left(y'(\xi)\right)^2\right]^{\frac{1}{2}} [y(\xi)]^{-\frac{1}{2}}\, d\xi \tag{3.83}$$

where ξ is a dummy variable of integration. If

$$f_0(x, y, y') = \frac{1}{(2g)^{\frac{1}{2}}} \left[1 + \left(y'(\xi)\right)^2\right]^{\frac{1}{2}} [y(\xi)]^{-\frac{1}{2}} \tag{3.84}$$

denotes the integrand, its partial derivatives are

$$\begin{aligned}
\frac{\partial f_0}{\partial y} &= -\left(\frac{1}{2\sqrt{g}}\right) \frac{1}{2y} \sqrt{\frac{1 + (y')^2}{y}} \\
\frac{\partial f_0}{\partial y'} &= \left(\frac{1}{2\sqrt{g}}\right) \frac{y'}{\sqrt{y(1 + (y')^2)}} \\
\frac{\partial f_0}{\partial x} &= 0
\end{aligned}$$

Therefore, the Euler-Lagrange equation for this problem is

$$\frac{\partial f_0}{\partial y} - \frac{d}{dx}\left[\frac{\partial f_0}{\partial y'}\right] = \left(\frac{1}{2\sqrt{g}}\right) \left\{ -\frac{1}{2y} \sqrt{\frac{1 + (y')^2}{y}} - \frac{d}{dx}\left[\frac{y'}{\sqrt{y(1 + (y')^2)}}\right] \right\} = 0$$

Hence, we wish to solve the following differential equation with boundary conditions for the shortest path from $(x_A, y_A) = (0, 0)$ to (x_B, y_B):

$$-\frac{1}{2y} \sqrt{\frac{1 + (y')^2}{y}} - \frac{d}{dx}\left[\frac{y'}{\sqrt{y(1 + (y')^2)}}\right] = 0 \tag{3.85}$$

$$y(0) = 0 \tag{3.86}$$

$$y(x_B) = y_B \tag{3.87}$$

Note that (3.85), (3.86), and (3.87) form a two-point boundary-value problem. The elementary theory of ordinary differential equations does not prepare one for solving a two-point boundary-value problem. We will discuss two-point boundary-value problems in more detail when we turn our attention to optimal control theory. For our present discussion of the brachistochrone problem, we will attempt a solution by first noting that

$$\frac{\partial f_0}{\partial x} = 0 \tag{3.88}$$

which suggests that we make use of the Beltrami identity (3.75) of Section 3.1.9 to assert

$$y' \frac{\partial f_0}{\partial y'} - f_0 = K_0 \text{ , a constant} \tag{3.89}$$

Therefore

$$\begin{aligned} 2\sqrt{g}K_0 &= y' \left[\frac{y'}{\sqrt{y(1+(y')^2)}} \right] - \sqrt{\frac{1+(y')^2}{y}} \\ &= \frac{(y')^2 - \left(1+(y')^2\right)}{\sqrt{y(1+(y')^2)}} \\ &= \frac{-1}{\sqrt{y(1+(y')^2)}} \end{aligned} \tag{3.90}$$

By introducing a new constant, (3.90) may be put in the form

$$\sqrt{y(1+(y')^2)} = K \equiv \frac{-1}{2\sqrt{g}K_0} \tag{3.91}$$

Thus

$$\frac{dy}{dx} = \sqrt{\frac{K^2 - y}{y}}$$

from which we obtain

$$dx = dy \sqrt{\frac{y}{K^2 - y}} \tag{3.92}$$

It may be shown that the parametric solution

$$\begin{aligned} x(t) &= \frac{K^2}{2}(t - \sin t) \\ y(t) &= \frac{K^2}{2}(1 - \cos t) \end{aligned}$$

satisfies (3.92). The constant K is determined by the boundary conditions. If we consider the case of $(x_B, y_B) = (\pi - 2, 2)$, then the brachistochrone path becomes

$$\begin{aligned} x(t) &= 2(t - \sin t) \\ y(t) &= 2(1 - \cos t) \end{aligned}$$

This path is shown by the dotted line in Figure 3.1.

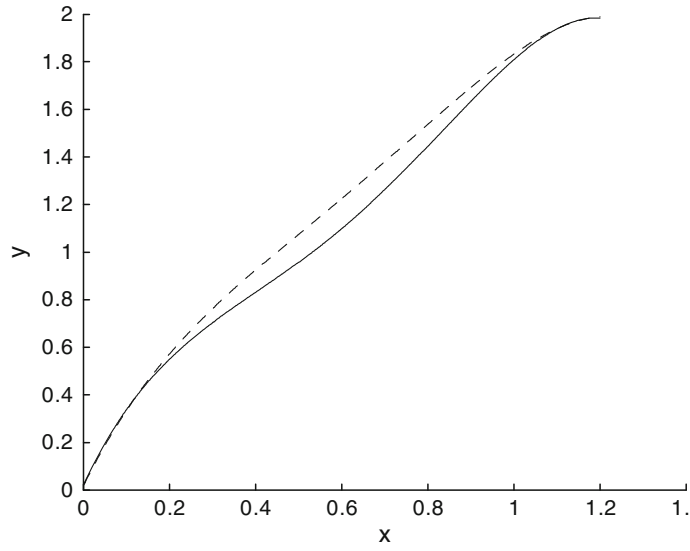


Fig. 3.1 An Approximate solution of the brachistochrone problem

It is instructive to also solve this problem numerically by exploiting off-the-shelf nonlinear programming software. In particular, by using the definition of a Riemann integral, the criterion (3.83) may be given the discrete approximation

$$\min J\,(x_0,\ldots,x_N,y_0,\ldots,y_N) = \frac{1}{(2g)^{\frac{1}{2}}}\sum_{i=1}^{N}\left[1+\left(\frac{y_i-y_{i-1}}{\Delta x}\right)^2\right]^{\frac{1}{2}}[y_i]^{-\frac{1}{2}}\,\Delta x \tag{3.93}$$

$$(x_0,y_0) = (x_A,y_A) = (0,0) \tag{3.94}$$

$$(x_N,y_N) = (x_B,y_B) = (\pi-2,2) \tag{3.95}$$

The Optimization Toolbox of MATLAB may be employed to solve the finite-dimensional mathematical program (3.93), (3.94), and (3.95) to determine an approximate trajectory, shown by the solid line in Figure 3.1 when $N = 11$. Note that the approximate solution compares favorably with the exact solution (the dashed line). Better agreeement may be achieved by increasing N.

3.3 Continuous-Time Optimal Control

In the theory of optimal control we are concerned with extremizing (maximizing or minimizing) a criterion functional subject to constraints. Both the criterion and the constraints are articulated in terms of two types of variables: control variables and

state variables. The state variables obey a system of first-order ordinary differential equations whose righthand sides typically depend on the control variables; initial values of the state variables are either specified or meant to be determined in the process of solving a given optimal control problem. Consequently, when the control variables and the state initial conditions are known, the state dynamics may be integrated and the state trajectories found. In this sense, the state variables are not really the decision varibles; rather, the control variables are the fundamental decision variables.

For reasons that will become clear, we do not require the control variables to be continuous; instead we allow the control variables to exhibit jump discontinuities. Furthermore, the constraints of an optimal control problem may include, in addition to the state equations and state initial conditions already mentioned, constraints expressed purely in terms of the controls, constraints expressed purely in terms of the state variables, and constraints that involve both control variables and state variables. The set of piecewise continuous controls satisfying the constraints imposed on the controls is called the set of *admissible controls*. Thus, the admissible controls are roughly analogous to the feasible solutions of a mathematical program.

Consider now the following canonical form of the continuous-time optimal control problem with pure control constraints:

$$\text{criterion} : \min J\left[x\left(t\right), u\left(t\right)\right] = K\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} f_0\left[x\left(t\right), u\left(t\right), t\right] dt \tag{3.96}$$

subject to the following:

$$\text{state dynamics} : \frac{dx}{dt} = f\left(x\left(t\right), u\left(t\right), t\right) \tag{3.97}$$

$$\text{initial conditions} : x\left(t_0\right) = x_0 \in \Re^m \qquad t_0 \in \Re^1 \tag{3.98}$$

$$\text{terminal conditions} : \Psi\left[x\left(t_f\right), t_f\right] = 0 \qquad t_f \in \Re^1 \tag{3.99}$$

$$\text{control constraints} : u\left(t\right) \in U \quad \forall t \in \left[t_0, t_f\right] \tag{3.100}$$

where for each instant of time $t \in \left[t_0, t_f\right] \subset \Re_+^1$:

$$x\left(t\right) = \left(x_1\left(t\right), x_2\left(t\right), \ldots, x_n\left(t\right)\right)^T \tag{3.101}$$

$$u\left(t\right) = \left(u_1\left(t\right), u_2\left(t\right), \ldots, u_m\left(t\right)\right)^T \tag{3.102}$$

$$f_0 : \Re^n \times \Re^m \times \Re^1 \longrightarrow \Re^1 \tag{3.103}$$

$$f : \Re^n \times \Re^m \times \Re^1 \longrightarrow \Re^n \tag{3.104}$$

$$K : \Re^n \times \Re^1 \longrightarrow \Re^1 \tag{3.105}$$

$$\Psi : \Re^n \times \Re^1 \longrightarrow \Re^r \tag{3.106}$$

We will use the notation $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ to refer to the above canonical optimal control problem. We assume the functions $f_0(.,.,)$, $\Psi(.,.)$, $K(.,.)$, and $f(.,.,.)$ are everywhere once continuously differentiable with respect to their arguments. In fact, we employ the following definition:

Definition 3.2. *Regularity for* $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$. *We shall say optimal control problem* $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ *defined by (3.96), (3.97), (3.98), (3.99), and (3.100) is regular provided* $f(x, u, .)$, $f_0(x, u, .)$, $\Psi[x(t_f), t_f]$, *and* $K[x(t_f), t_f]$ *are everywhere once continuously differentiable with respect to their arguments.*

We also formally define the notion of an admissible solution for

$$OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$$

Definition 3.3. *Admissible control trajectory. We say that the control trajectory* $u(t)$ *is admisible relative to* $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ *if it is piecewise continuous for all time* $t \in [t_0, t_f]$ *and* $u \in U$.

Note that the initial time and the terminal time may be unknowns in the continuous-time optimal control problem. Moreover, the initial values $x(t_0)$ and final values $x(t_f)$ may be unknowns. Of course, the initial and/or final values may also be stipulated. The unkowns are the state variables x and the control variables u. It is critically important to realize that the state variables will generally be completely determined when the controls and initial states are known. Consequently, the "true" unknowns are the control variables u. Note also that we have not been specific about the vector space to which

$$x = (x(t) : t \in [t_0, t_f])$$
$$u = (u(t) : t \in [t_0, t_f])$$

belong. This is by design, as we shall initially discuss the continuous-time optimal control problem by developing intuitive dynamic extensions of the notion of stationarity and an associated calculus for variations of $x(t)$ and $u(t)$. In the next chapter, we shall introduce results from the theory of infinite-dimensional mathematical programming that allow a more rigorous mathematical analysis of the continuous-time optimal control problem.

3.3.1 Necessary Conditions for Continuous-Time Optimal Control

We begin by commenting that we employ the notation $\delta\mathcal{L}$, which is used in classical and traditional references on the theory of optimal control, for what we have defined in Section 3.1.2 to be the variation of the functional $\mathcal{L}$. Relying as it does on the variational notation introduced early in Section 3.1.2, our derivation of the optimal

control necessary conditions in this section will be informal. To derive necessary conditions in such a manner, we will need the *variation of the state vector* x, denoted by δx. We will make use of the relationship

$$dx = \delta x + \dot{x} dt \tag{3.107}$$

that identifies δx, the variation of x, as that part of the total change dx not attributable to time. Variations of other entities, such as u, are denoted in a completely analogous fashion. We start our derivation of optimal control necessary conditions by pricing out all constraints to obtain the Lagrangean

$$\mathcal{L} = K\left[x\left(t_f\right), t_f\right] + \nu^T \Psi\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \left\{f_0\left(x, u, t\right) + \lambda^T \left[f\left(x, u, t\right) - \dot{x}\right]\right\} dt \tag{3.108}$$

Using the variational calculus chain rule developed earlier, we may state the variation of the Lagrangean $\mathcal{L}$ as

$$\begin{aligned} \delta\mathcal{L} = &\left[\Phi_t\left(t_f\right) dt_f + \Phi_x\left(t_f\right) dx\left(t_f\right) + f_0\left(t_f\right) dt_f\right] - f_0\left(t_0\right) dt_0 \\ &+ \int_{t_0}^{t_f} \left[H_x \delta x + H_u \delta u - \lambda^T \delta \dot{x}\right] dt \end{aligned} \tag{3.109}$$

where

$$H\left(x, u, \lambda, t\right) \equiv f_0\left(x, u, t\right) + \lambda^T f\left(x, u, t\right) \tag{3.110}$$

is the *Hamiltonian* and

$$\Phi\left(t_f\right) \equiv K\left[x\left(t_f\right), t_f\right] + \nu^T \Psi\left[x\left(t_f\right), t_f\right] \tag{3.111}$$

$$f_0\left(t_0\right) \equiv f_0\left[x\left(t_0\right), u\left(t_0\right), t_0\right] \tag{3.112}$$

$$f_0\left(t_f\right) \equiv f_0\left[x\left(t_f\right), u\left(t_f\right), t_f\right] \tag{3.113}$$

$$\Phi_t \equiv \frac{\partial \Phi}{\partial t} \qquad \Phi_x \equiv \frac{\partial \Phi}{\partial x} \tag{3.114}$$

$$f_{0x} \equiv \frac{\partial f_0}{\partial x} \qquad f_x \equiv \frac{\partial f}{\partial x} \tag{3.115}$$

$$H_x \equiv \frac{\partial H}{\partial x} = \left(\nabla_x H\right)^T \qquad H_u \equiv \frac{\partial H}{\partial u} = \left(\nabla_u H\right)^T \tag{3.116}$$

We next turn our attention to the term

$$\begin{aligned} I &\equiv \int_{t_0}^{t_f} \left(-\lambda^T \delta \dot{x}\right) dt = -\int_{t_0}^{t_f} \lambda^T \frac{d}{dt}\left(\delta x\right) dt \\ &= -\int_{t_0}^{t_f} \lambda^T d\left(\delta x\right) \end{aligned}$$

appearing in (3.109). In particular, using the rule for integrating by parts[1] this integral becomes

$$\begin{aligned} I &= \lambda^T(t_0)\,\delta x(t_0) - \lambda^T(t_f)\,\delta x(t_f) + \int_{t_0}^{t_f} \left(d\lambda^T\right)\delta x \\ &= \lambda^T(t_0)\,\delta x(t_0) - \lambda^T(t_f)\,\delta x(t_f) + \int_{t_0}^{t_f} \left(\frac{d\lambda^T}{dt}\delta x\right) dt \end{aligned} \tag{3.117}$$

We also note that

$$\delta x(t_f) = dx(t_f) - \dot{x}(t_f)\,dt_f \tag{3.118}$$

$$\delta x(t_0) = dx(t_0) - \dot{x}(t_0)\,dt_0 \tag{3.119}$$

from the definition of a variation of the state vector. Using (3.117) in (3.109) gives

$$\begin{aligned} \delta\mathcal{L} = &\left[\Phi_t(t_f)\,dt_f + \Phi_x^T(t_f)\,dx(t_f) + f_0(t_f)\,dt_f\right] - f_0(t_0)\,dt_0 \\ &+ \int_{t_0}^{t_f} [H_x\delta x + H_u\delta u]\,dt \\ &+ \lambda^T(t_0)\,\delta x(t_0) - \lambda^T(t_f)\,\delta x(t_f) + \int_{t_0}^{t_f}\left(\frac{d\lambda^T}{dt}\delta x\right)dt \end{aligned} \tag{3.120}$$

Using (3.118) and (3.119) in (3.120) gives

$$\begin{aligned} \delta\mathcal{L} = &\left[\Phi_t(t_f)\,dt_f + \Phi_x(t_f)\,dx(t_f) + f_0(t_f)\,dt_f\right] - f_0(t_0)\,dt_0 \\ &+ \lambda^T(t_0)\left[dx(t_0) - \dot{x}(t_0)\,dt_0\right] - \lambda^T(t_f)\left[dx(t_f) - \dot{x}(t_f)\,dt_f\right] \\ &+ \int_{t_0}^{t_f}\left[\left(H_x + \frac{d\lambda^T}{dt}\right)\delta x + H_u\delta u\right]dt \end{aligned} \tag{3.121}$$

It follows from (3.121), upon rearranging and collecting terms, that

$$\begin{aligned} \delta\mathcal{L} = &\left[\Phi_t(t_f) + f_0(t_f) + \lambda^T(t_f)\,\dot{x}(t_f)\right]dt_f \\ &+ \left[\Phi_x^T(t_f) - \lambda^T(t_f)\right]dx(t_f) + \lambda^T(t_0)\,dx(t_0) \\ &- \left[f_0(t_0) + \lambda^T(t_0)\,\dot{x}(t_0)\right]dt_0 \\ &+ \int_{t_0}^{t_f}\left[\left(H_x + \dot{\lambda}^T\right)\delta x + H_u\delta u\right]dt \end{aligned} \tag{3.122}$$

[1] Integration by parts: $\int u dv = uv - \int v du$.

We see from (3.122) that, in order for $\delta\mathcal{L}$ to vanish for arbitrary admissible variations, the coefficient of each individual differential and variation must be zero. That is, for the case of no explicit control constraints, $\delta\mathcal{L} = 0$ is ensured by the following necessary conditions for optimality:

1. state dynamics:

$$\frac{dx}{dt} = f\left(x\left(t\right), u\left(t\right), t\right) \tag{3.123}$$

2. initial time conditions:

$$H\left(t_0\right) = 0 \text{ and } \lambda\left(t_0\right) = 0 \Longrightarrow f_0\left[x\left(t_0\right), u\left(t_0\right), t_0\right] = 0 \tag{3.124}$$

$$x\left(t_0\right) = x_0 \in \Re^m \tag{3.125}$$

3. adjoint equations:

$$\dot{\lambda} = -H_x = -f_{0x} - \lambda^T f_x \tag{3.126}$$

4. transversality conditions:

$$\lambda\left(t_f\right) = \Phi_x\left(t_f\right) = K_x\left[x\left(t_f\right), t_f\right] + \nu^T \Psi_x\left[x\left(t_f\right), t_f\right] \tag{3.127}$$

5. terminal time conditions:

$$\Psi_t\left[x\left(t_f\right), t_f\right] = 0 \tag{3.128}$$

$$-H\left(t_f\right) = \Phi_t\left(t_f\right) \tag{3.129}$$

where

$$H\left(t_f\right) \equiv f_0\left[x\left(t_f\right), u\left(t_f\right), t_f\right] + \lambda^T\left(t_f\right) f\left[x\left(t_f\right), u\left(t_f\right), t_f\right]$$
$$\Phi_t\left(t_f\right) \equiv K_t\left[x\left(t_f\right), t_f\right] + \nu^T \Psi_t\left[x\left(t_f\right), t_f\right]$$

6. minimum principle:

$$H_u\left(x, u, \lambda, t\right) = 0 \tag{3.130}$$

Note carefully that a two-point boundary-value problem is an explicit part of these necessary conditions. That is to say, we need to solve a system of ordinary differential equations, namely the original state dynamics (3.123) together with the adjoint equations (3.126), given the initial values of the state variables (3.125) and the transversality conditions (3.127) imposed on the adjoint variables at the terminal time; this will typically be the case even when the initial time t_0, the terminal time t_f and the initial state $x\left(t_0\right)$ are fixed and the terminal state $x\left(t_f\right)$ is free. Note also that when the initial time t_0 is fixed, we do not enforce (3.124), since dt_0 will vanish.

To develop necessary conditions for the case of explicit control constraints, we invoke an intuitive argument. In particular, we argue that the total variation

expressed by $\delta\mathcal{L}$ must be nonnegative if the current solution is optimal; otherwise, there would exist a potential to decrease $\mathcal{L}$ (and hence J) and such a potential would not be consistent with having achieved a minimum. Since it is only the variation δu that is impacted by the constraints $u \in U$ and which can no longer be arbitrary, we may invoke all the conditions developed above except the one requiring the coefficient of δu to vanish; instead, we require

$$\delta\mathcal{L} = \int_{t_0}^{t_f} (H_u \delta u)\, dt \geq 0 \tag{3.131}$$

In order for condition (3.131) to be satisfied for all admissible variations δu, we require

$$H_u \delta u = H_u \left(u - u^*\right) \geq 0 \quad \forall u \in U \tag{3.132}$$

where we have expressed the variation of u as

$$\delta u = u - u^*$$

which describes feasible directions rooted at the optimal control solution $u^* \in U$ when the set U is convex. Inequality (3.132) is the correct form of the minimum principle when there are explicit, pure control constraints forming a convex set U; it is known as a *variational inequality*. We will have much more to say about variational inequalities when we employ functional analysis to study the continuous-time optimal control problem in the next chapter.

The above discussion has been a constructive proof of the following result:

Theorem 3.6. *Necessary conditions for continuous-time optimal control problem. When the variations of x, $\dot{x}$ and u are well defined and linear in their increments, the set of feasible controls U is convex, and regularity in the sense of Definition 3.2 obtains, the conditions (3.123), (3.124), (3.125), (3.126), (3.127), (3.128), (3.129), and (3.132) are necessary conditions for a solution of the optimal control problem $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ defined by (3.96), (3.97), (3.98), (3.99), and (3.100).*

3.3.2 Necessary Conditions with Fixed Terminal Time, No Terminal Cost, and No Terminal Constraints

A frequently encountered problem type is

$$\min\ J = \int_{t_0}^{t_f} f_0\,(x, u, t)\, dt$$

subject to

$$\frac{dx}{dt} = f\,(x, u, t)$$

$$x\left(t_0\right) = x_0$$

where both t_0 and t_f are fixed; also x_0 is fixed. In this case

$$\delta\mathcal{L} = \int_{t_0}^{t_f} \left[H_x \delta x + H_u \delta u - \lambda^T \delta\dot{x}\right] dt \tag{3.133}$$

Using integration by parts, we have

$$\begin{aligned}\int_{t_0}^{t_f} \left(-\lambda^T \delta\dot{x}\right) dt &= -\int_{t_0}^{t_f} \lambda^T d\left(\delta x\right) \\ &= \lambda^T\left(t_0\right)\delta x\left(t_0\right) - \lambda^T\left(t_f\right)\delta x\left(t_f\right) + \int_{t_0}^{t_f} \left(\frac{d\lambda^T}{dt}\delta x\right) dt \\ &= -\lambda^T\left(t_f\right)\delta x\left(t_f\right) + \int_{t_0}^{t_f} \left(\frac{d\lambda^T}{dt}\delta x\right) dt \end{aligned} \tag{3.134}$$

We also know that

$$\delta x\left(t_f\right) = dx\left(t_f\right) - \dot{x}\left(t_f\right) dt_f = dx\left(t_f\right) \tag{3.135}$$

so that (3.134) becomes

$$\int_{t_0}^{t_f} \left(-\lambda^T \delta\dot{x}\right) dt = -\lambda^T\left(t_f\right) dx\left(t_f\right) + \int_{t_0}^{t_f} \left(\frac{d\lambda^T}{dt}\delta x\right) dt$$

It follows that (3.133) becomes

$$\delta\mathcal{L} = -\lambda^T\left(t_f\right) dx\left(t_f\right) + \int_{t_0}^{t_f} \left[\left(H_x + \frac{d\lambda^T}{dt}\right)\delta x + H_u \delta u\right] dt \tag{3.136}$$

It is then immediate from (3.136) that $\delta\mathcal{L}$ vanishes when the following necessary conditions

$$H_x + \frac{d\lambda^T}{dt} = 0 \tag{3.137}$$

$$\lambda^T\left(t_f\right) = 0 \tag{3.138}$$

$$H_u = 0 \tag{3.139}$$

together with the original state dynamics, state initial condition and control constraints.

3.3.3 Necessary Conditions When the Terminal Time Is Free

In some applications the terminal time may not be fixed and so its variation will not be zero. We are interested in deriving necessary conditions for such problems and then in exploring how they may be used to solve an example problem. To illustrate how such conditions are derived, we consider the following simplified problem:

$$\min\ J = K\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} f_0\left(x, u, t\right) dt$$

subject to

$$\frac{dx}{dt} = f\left(x, u, t\right)$$
$$x\left(t_0\right) = x_0$$

where t_f is free, t_0 and x_0 are fixed, and

$$x_i\left(t_f\right) \text{ is fixed for } i = 1, \ldots, q < n$$
$$x_i\left(t_f\right) \text{ is free for } \ i = q+1, \ldots, n$$

We will denote the states that will be free at the terminal time by the following vector

$$x^{\text{free}} = \left(x_{q+1}, \ldots, x_n\right)^T$$

Moreover the terminal cost function is taken to be

$$K\left[x\left(t_f\right), t_f\right] = \phi\left[x^{\text{free}}(t_f), t_f\right]$$

to reflect that it depends only on those states that are free. We proceed as previously by pricing out the dynamics to create

$$\mathcal{L} = \phi\left[x^{\text{free}}(t_f), t_f\right] + \int_{t_0}^{t_f} \left\{f_0\left(x, u, t\right) + \lambda^T\left[f\left(x, u, t\right) - \dot{x}\right]\right\} dt$$

so that the variation of $\mathcal{L}$ is

$$\delta\mathcal{L} = \left(\frac{\partial\phi}{\partial t}dt + \frac{\partial\phi}{\partial x}dx\right)_{t=t_f} + \left(f_0\right)_{t=t_f} dt_f$$
$$+ \int_{t_0}^{t_f}\left[\left(\frac{\partial f_0}{\partial x} + \lambda^T\frac{\partial f}{\partial x}\right)\delta x + \left(\frac{\partial f_0}{\partial u} + \lambda^T\frac{\partial f}{\partial u}\right)\delta u - \lambda^T\delta\dot{x}\right] dt$$

Using the usual device of integrating by parts then collecting terms, the above expression becomes

$$\delta\mathcal{L} = \left[\left(\frac{\partial\phi}{\partial t} + f_0\right)dt_f + \frac{\partial\phi}{\partial x}dx\right]_{t=t_f} - \left[\lambda^T\delta x\right]_{t=t_f} + \left[\lambda^T\delta x\right]_{t=t_0} + \int_{t_0}^{t_f}\left[\left(\frac{\partial f_0}{\partial x} + \lambda^T\frac{\partial f}{\partial x} + \dot{\lambda}^T\right)\delta x + \left(\frac{\partial f_0}{\partial u} + \lambda^T\frac{\partial f}{\partial u}\right)\delta u\right]dt$$

We next observe $\delta x(t_0) = 0$ and exploit the identity

$$\delta x\left(t_f\right) = dx\left(t_f\right) - \dot{x}\left(t_f\right)dt_f$$

to rewrite $\delta\mathcal{L}$ as

$$\delta\mathcal{L} = \left[\left(\frac{\partial\phi}{\partial t} + H\right)dt_f + \left(\frac{\partial\phi}{\partial x} - \lambda^T\right)dx\right]_{t=t_f} + \int_{t_0}^{t_f}\left[\left(\frac{\partial H}{\partial x} + \dot{\lambda}^T\right)\delta x + \left(\frac{\partial H}{\partial u}\right)\delta u\right]dt$$

For the above we realize that

$$dx_i(t_f) = 0 \quad \text{for} \quad i = 1,\ldots,q < n \tag{3.140}$$

Therefore, $\delta\mathcal{L}$ will vanish when the following necessary conditions hold:

$$\frac{d\lambda_j}{dt} = -\frac{\partial H}{\partial x_j} \quad j = 1,\ldots,q$$

$$\lambda_j\left(t_f\right) = \begin{cases} \nu_j & j = 1,\ldots,q \\ \left(\dfrac{\partial\phi}{\partial x_j}\right)_{t=t_f} & j = q+1,\ldots,n \end{cases}$$

$$\frac{\partial H}{\partial u} = 0$$

$$0 = \left(\frac{\partial\phi}{\partial t} + H\right)_{t=t_f}$$

where the ν_j for $j = 1,\ldots,q$ are in effect additional control variables that must somehow be determined in order for the free terminal time problem we have posed to have a solution; their determination should result in stationarity of the Hamiltonian, in accordance with the minimum principle.

3.3.4 Necessary Conditions for Problems with Interior Point Constraints

Suppose there are interior boundary conditions

$$N\left[x\left(t_1\right), t_1\right] = 0 \tag{3.141}$$

where $t_0 < t_1 < t_f$ and $N : \Re^{n+1} \to \Re^q$. Constraints (3.141) are terminal constraints for the interval $[t_0, t_1]$. In this setting we take t_0, t_1, and t_f to be fixed; also $x\left(t_0\right)$ is fixed. We let t_1^- signify an instant in time just prior to t_1 and t_1^+ an instant just following t_1. We develop necessary conditions by adjoining (3.141) to the criterion so that

$$\begin{aligned} J_1 = {} & \Psi\left[x\left(t_f\right), t_f\right] + \pi^T N\left[x\left(t_1\right), t_1\right] \\ & + \int_{t_0}^{t_f} \left\{ f_0\left(x, u, t\right) + \lambda^T\left[f\left(x, u, t\right) - \dot{x}\right] \right\} dt \end{aligned}$$

Proceeding in the usual way we have

$$\begin{aligned} \delta J = {} & \left(\frac{\partial \Psi}{\partial t}\delta x\right)_{t=t_f} + \pi^T \frac{\partial N}{\partial t_1} dt_1 + \pi^T \frac{\partial N}{\partial x\left(t_1\right)} dx\left(t_1\right) \\ & - \left[\lambda^T \delta x\right]_{t_1^+}^{t_f} - \left[\lambda^T \delta x\right]_{t_0}^{t_1^-} + \left(H - \lambda^T \dot{x}\right)_{t=t_1^-} dt_1 - \left(H - \lambda^T \dot{x}\right)_{t=t_1^+} dt_1 \\ & + \int_{t_0}^{t_f} \left[\left(\dot{\lambda}^T + \frac{\partial H}{\partial x}\right)\delta x + \frac{\partial H}{\partial u}\delta u\right] dt \end{aligned}$$

We next employ the identities

$$dx\left(t_1\right) = \begin{cases} \delta x\left(t_1^-\right) + \dot{x}\left(t_1^-\right) dt_1 \\ \delta x\left(t_1^+\right) + \dot{x}\left(t_1^+\right) dt_1 \end{cases}$$

which lead, after some manipulation, to the following

$$\begin{aligned} \delta J = {} & \left[\left(\frac{\partial \Psi}{\partial t} - \lambda^T\right)\delta x\right]_{t=t_f} + \left[\lambda^T\left(t_1^+\right) - \lambda^T\left(t_1^-\right) + \pi^T \frac{\partial N}{\partial x\left(t_1\right)}\right] dx\left(t_1\right) \\ & + \left[H\left(t_1^-\right) - H\left(t_1^+\right) + \pi^T \frac{\partial N}{\partial t_1}\right] dt_1 + \left(\lambda^T \delta x\right)_{t=t_0} \\ & + \int_{t_0}^{t_f} \left[\left(\dot{\lambda}^T + \frac{\partial H}{\partial x}\right)\delta x + \frac{\partial H}{\partial u}\delta u\right] dt \end{aligned}$$

Obviously $\delta x(t_0)$ vanishes, given $x(t_0)$ is fixed. Our task is to select $\lambda\left(t_1^-\right)$ and $H\left(t_1^-\right)$ to cause the coefficients of $dx(t_1)$ and dt_1 to vanish. Doing so yields

$$\lambda^T\left(t_1^-\right) = \lambda^T\left(t_1^+\right) + \pi^T \frac{\partial N}{\partial x(t_1)} \tag{3.142}$$

$$H\left(t_1^-\right) = H\left(t_1^+\right) - \pi^T \frac{\partial N}{\partial t_1} \tag{3.143}$$

We of course also have the traditional necessary conditions

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} \tag{3.144}$$

$$\lambda^T\left(t_f\right) = \left(\frac{\partial \Psi}{\partial x}\right)_{t=t_f} \tag{3.145}$$

$$\frac{\partial H}{\partial u} = 0 \tag{3.146}$$

3.3.5 Dynamic Programming and Optimal Control

Another profound contribution to the field of dynamic optimization in the twentieth century was the theory and computational paradigm known as *dynamic programming*, frequently referred to as "DP." Dynamic programming, developed by the reknown American mathematician Richard Bellman, provides an alternative approach to the study and solution of optimal control problems for both discrete and continuous-time. It is important, however, to recognize that dynamic programming is much more general than optimal control theory in that it does not require that the dynamic optimization problem of interest be a calculus of variations problem.

The fundamental result that provides the foundation for dynamic programming is known as the *Principle of Optimality*, which Bellman (1957) originally stated as

> An optimal policy has the property that, whatever the initial state and decision [*control* in our language] are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

This principle accords with common sense and is quite easy to prove by contradiction. From the principle of optimality (POO) we directly obtain the fundamental recursive relationship employed in dynamic programming.

Let us define

$$J^*(x,t) \tag{3.147}$$

to be the *optimal performance function* (OPF) for the continuous-time optimal control problem introduced previously. The OPF is the minimized value of the objective functional of the continuous-time optimal control problem. Note carefully that (3.147) is referred to as a *function* and not as a *functional* because it is viewed

as the specific value of the performance functional corresponding to an optimal process starting at state x at time t. This distinction is critical to avoiding misuse and misinterpretation of the results we next develop. According to the POO, if $J^*(x,t)$ is the OPF for a problem starting at state x at time t, then it must be that $J^*(x+\Delta x, t+\Delta t)$ is the OPF for that portion of the optimal trajectory starting at state $x+\Delta x$ at time $t+\Delta t$. However, during the interval of time $[t, t+\Delta t]$, the only change in the OPF is that due to the integrand $f_0(x,u,t)$ of the objective functional acting for Δt units of time with effect $f_0(x,u,t)\,\Delta t$. It, therefore, follows from the POO that the OPF values for (x,t) and $(x+\Delta x, t+\Delta t)$ are related according to

$$J^*(x,t) = \min_{u\in U}\left[f_0(x,u,t)\,\Delta t + J^*(x+\Delta x, t+\Delta t)\right] \tag{3.148}$$

Our development of (3.148), as well as of subsequent results in this section, depends on two mild but important regularity conditions, namely

1. $J^*(x,t)$ is single valued; and
2. $J^*(x,t)$ is $\mathcal{C}^1$(continuously differentiable).

This means in effect that solutions to the problem of constrained minimization of the objective functional vary continuously with respect to the initial conditions.

Because of the aforementioned regularity assumptions, we may make a Taylor series expansion of the OPF $J^*(x+\Delta x, t+\Delta t)$ about the point (x,t). That expansion takes the form

$$J^*(x+\Delta x, t+\Delta t) = J^*(x,t) + \left[\nabla_x J^*(x,t)\right]^T \Delta x + \frac{\partial J^*(x,t)}{\partial t}\Delta t + \ldots \tag{3.149}$$

where $[\nabla_x J^*(x,t)]^T$ is the row vector

$$\left[\nabla_x J^*(x,t)\right]^T = \left[\frac{\partial J^*(x,t)}{\partial x_1}, \frac{\partial J^*(x,t)}{\partial x_2}, \ldots, \frac{\partial J^*(x,t)}{\partial x_n}\right] \equiv \frac{\partial J^*(x,t)}{\partial x}$$

Inserting (3.149) in (3.148), dividing by Δt and retaining only the linear terms, we obtain

$$0 = \min_{u\in U}\left[f_0(x,u,t) + \frac{\partial J^*(x,t)}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial J^*(x,t)}{\partial t}\right]$$

Taking the limit of this last expression as $\Delta t \longrightarrow 0$ yields

$$(-1)\frac{\partial J^*(x,t)}{\partial t} = \min_{u\in U}\left[f_0(x,u,t) + \frac{\partial J^*(x,t)}{\partial x} f(x,u,t)\right] \tag{3.150}$$

since $\dot{x} = f(x,u,t)$. Result (3.150) is the basic recursive relationship of dynamic programming and is known as *Bellman's equation*.

Before continuing with our analysis of the relationship between the continuous-time optimal control problem and dynamic programming, we must develop a deeper

understanding of the adjoint variables. To this end, we observe that the Lagrangean for an arbitrary initial time t_0 will be

$$\mathcal{L}(x,u,\nu,\lambda,t) = K[x(T),T] + \nu^T \Psi\left[x\left(t_f\right),t_f\right] + \int_{t_0}^{t_f} \left\{ f_0(x,u,t) + \lambda^T [f(x,u,t) - \dot{x}] \right\} dt \quad (3.151)$$

for which the key constraints have been priced out and added to the objective functional. For the analysis that follows, we shall consider, unless otherwise stated, that we are on an optimal solution trajectory. On that optimal trajectory we take our criterion to be $\mathcal{L}(x,u,\nu,\lambda,t) = J[x(t_0),t_0]$. Integrating (3.151) by parts we obtain

$$\begin{aligned} J[x(t_0),t_0] = {} & K\left[x\left(t_f\right),t_f\right] + \nu^T \Psi\left[x\left(t_f\right),t_f\right] \\ & + \int_{t_0}^{t_f} \left[H(x,u,\lambda,t) + \left(\dot{\lambda}\right)^T x \right] dt \\ & - \left[\lambda^T\left(t_f\right) x\left(t_f\right) - \lambda^T(t_0)\, x(t_0) \right] \end{aligned} \quad (3.152)$$

upon using the definition of the Hamiltonian

$$H(x,u,\lambda,t) = f_0(x,u,t) + \lambda^T f(x,u,t)$$

By inspection of (3.152), it is apparent that partial differentiation of the criterion with respect to the initial state along an optimal solution trajectory yields

$$\frac{\partial J[x(t_0),t_0]}{\partial x(t_0)} = \left[\lambda^*(t_0)\right]^T \quad (3.153)$$

Since time t_0 in this analysis is arbitrary so that $\left[t_0,t_f\right]$ may correspond to any portion of the optimal trajectory, we see from expression (3.153) that the adjoint variable measures the sensitivity of the OPF to changes in the state variable at the start of the time interval under consideration. That is, the adjoint variables are dynamic dual variables for the state dynamics, and we may more generally write

$$\frac{\partial J^*}{\partial x} = \lambda^{*T} \quad (3.154)$$

Result (3.153) means that Bellman's equation (3.150) may be restated as

$$\begin{aligned} (-1)\frac{\partial J(x,t)}{\partial t} &= \min_{u\in U}\left[f_0(x,u,t) + \frac{\partial J(x,t)}{\partial x} f(x,u,t) \right] \\ &= \min_{u\in U}\left[H\left(x,u,\frac{\partial J(x,t)}{\partial x},t\right)\right] \equiv H^0\left(x,\frac{\partial J(x,t)}{\partial x},t\right) \end{aligned}$$

which can in turn be restated as

$$H^0\left(x, \frac{\partial J(x,t)}{\partial x}, t\right) + \frac{\partial J(x,t)}{\partial t} = 0 \tag{3.155}$$

where H^0 is the Hamiltonian after the optimal control law obtained from the minimum principle is used to eliminate control variables. The partial differential equation (3.155) is known as the Hamilton-Jacobi equation (HJE). The appropriate boundary condition for (3.155) comes from recognizing that at the terminal time the OPF must equal the terminal value. Hence

$$J\left[x\left(t_f\right), t_f\right] = K\left[x\left(t_f\right), t_f\right] + \nu^T \Psi\left[x\left(t_f\right), t_f\right] \tag{3.156}$$

is the appropriate boundary condition.

Our use of the dynamic programming concept of an optimal performance function (OPF) in this section has shown that the adjoint variables are true dynamic dual variables expressing the sensitivity of the OPF to changes in the initial state variable. We have also seen that the continuous-time optimal control problem has a necessary condition known as the Hamilton-Jacobi (partial differential) equation. Although sufficiency (of the HJE representation of the continuous-time optimal control problem) can be established under certain regularity conditions, we do not pursue that result here, partly because of dynamic programming's (and the HJE's) so-called "curse of dimensionality." This memorable phrase refers to the fact that, in order to use dynamic programming for a general problem, we must employ a grid of points for every component of $x(t) \in \Re^n$ to approximate the OPF by interpolation on that grid. Let us assume, for the sake of discussion, that we use ten (10) grid points for each component of $x(t) \in \Re^n$. The result is that the number of grid points (number of values of the OPF) to be stored is 10^n for each instant of time considered!

3.3.6 *Second-Order Variations in Optimal Control*

Sometimes additional information is needed beyond the necessary conditions derived above and based on the first-order variation. These additional conditions depend on second-order variations. To derive them, we consider the problem

$$\text{criterion}: \min J[x(t), u(t)] = K\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} f_0[x(t), u(t), t]\, dt \tag{3.157}$$

subject to

$$\text{state dynamics}: \frac{dx}{dt} = f(x(t), u(t), t) \tag{3.158}$$

$$\text{initial conditions}: x(t_0) = x_0 \in \Re^m \qquad t_0 \in \Re^1 \tag{3.159}$$

$$\text{terminal conditions} : \Psi\left[x\left(t_f\right), t_f\right] = 0 \qquad t_f \in \Re^1 \tag{3.160}$$

$$\text{control constraints} : u(t) \in \Re^m \quad \forall t \in \left[t_0, t_f\right] \tag{3.161}$$

Note we have made the simplifying assumption that there are no control constraints.

Our prior derivation of the Hamilton-Jacobi equation (HJE) tells us that

$$H^*\left(x, \frac{\partial J^*(x,t)}{\partial x}, t\right) + \frac{\partial J^*(x,t)}{\partial t} = 0 \tag{3.162}$$

where

$$H^*\left(x, \frac{\partial J^*(x,t)}{\partial x}, t\right) = \min_u \left[H\left(x, u, \frac{\partial J^*(x,t)}{\partial x}, t\right)\right] \tag{3.163}$$

Since minimization on the righthand side of (3.163) is unconstrained, the necessary conditions for a finite-dimensional unconstrained local minimum developed in Chapter 2 apply; that is

$$\frac{\partial H}{\partial u} = 0 \tag{3.164}$$

$$\frac{\partial^2 H}{\partial u^2} \geq 0 \tag{3.165}$$

for all $t \in \left[t_0, t_f\right]$. For (3.163), we are able to use the conditions for finite-dimensional mathematical programs because (3.163) is meant to hold separately at each instant of time. The conditions (3.164) and (3.165) are recognized as necessary conditions for an unconstrained local minimum of (3.163). Inequality (3.165) is called the *Legendre-Clebsch condition*.

As in prior discussions, we now form the Lagrangean

$$\mathcal{L} = K\left[x(T), T\right] + \nu^T \Psi\left[x(T), T\right] + \int_{t_0}^{t_f} \left\{f_0(x,u,t) + \lambda^T\left[f(x,u,t) - \dot{x}\right]\right\} dt \tag{3.166}$$

where our notation is identical to that introduced previously. We consider small perturbations from the extremal path corresponding to the minimization of $\mathcal{L}$; these small changes are a result of small perturbations $\delta x(t_0)$ of the initial state $x(t_0)$. We of course denote the variations of state, adjoint and control variables corresponding to these perturbations by $\delta x(t)$, $\delta\lambda(t)$, and $\delta u(t)$, respectively. We let

$$F(x, \dot{x}, u, \lambda, \dot{\lambda}) = 0$$

be an abstract representation of the system of equations resulting from linearizing the state dynamics, the adjoint equations and the minimum principle $\partial H/\partial u = 0$. It can be shown that

$$\delta F(x, \dot{x}, u, \lambda, \dot{\lambda}) = 0$$

is assured by the following

$$\delta \dot{x} = f_x \delta x + f_u \delta u \tag{3.167}$$

$$\delta \dot{\lambda} = -(H_{xx}\delta x)^T - (\delta\lambda)\, f^T - (H_{xu}\delta u)^T \tag{3.168}$$

$$\begin{aligned} \delta H_u &= (H_{ux}\delta x)^T + (\delta\lambda)\, H_{u\lambda}^T + (H_{uu}\delta u)^T \\ &= (H_{ux}\delta x)^T + (\delta\lambda)\, f_u + (H_{uu}\delta u)^T = 0 \end{aligned} \tag{3.169}$$

Because $x(t_0) = x_0$ and $\lambda(t_f)$ is specified by the transversality conditions, the system (3.167), (3.168), and (3.169) constitutes a two-point boundary-value problem provided (3.169) may be solved to obtain an expression for the variation δu; this requires that the Hessian H_{uu} be a *nonsingular matrix*. That is, when H_{uu} is invertible, we may completely characterize optimal solutions through the second variation equations (3.167), (3.168), and (3.169). We, therefore, call an optimal control corresponding to $H_{uu} = 0$ a *singular* control. Moreover, when $H_u = 0$, it must be that $H_{uu} = 0$ is *singular*, preventing the necessary conditions (3.164) and (3.165) for the minimum principle from yielding information regarding the optimal control. The interested reader may easily extend the above results on singular controls in the absence of constraints to the case of explicit control constaints and is encouraged to do so as a training device.

3.3.7 Singular Controls

As we have noted above, a simple definition of singular controls is that they are controls that arise when

$$\frac{\partial^2 H}{\partial u_i^2} = 0 \text{ for all } i = 1, 2, \ldots, m \tag{3.170}$$

Singular controls cannot be found from the unembellished minimum principle; rather they must be found from information that supplements the minimum principle. That information is obtained in any way that is consistent with the conditions that lead to singularity and can be thought of as invoking additional necessary conditions beyond the usual first-order and second-order conditions, which are trivially satisfied.

For problems that are linear in the controls u, the coefficient of u in the Hamiltonian is always $H_u = 0$, a circumstance that guarantees (3.170) is fulfilled.

One can seek functional equations describing the singular control of interest by forming differential equations based on successive time derivatives of H_u; that is

$$\frac{d^k}{dt^k}(H_u) = 0 \quad k = 1, 2, \ldots \tag{3.171}$$

In other cases, *Tait's necessary condition* may be employed:

$$(-1)^n \frac{\partial}{\partial u}\left[\left(\frac{d}{dt}\right)^{2n} H_u\right] \geq 0 \tag{3.172}$$

See Tait (1965) for a detailed presentation and proof of this result.

3.3.8 Sufficiency in Optimal Control

The necessary conditions considered in this chapter may only be used to find a globally optimal solution if we are able to uncover and compare *all* of the solutions of them. This is of course not in general possible for mathematical programming, variational and optimal control problems. Consequently, we are interested in this section in regularity conditions that make the optimal control necessary conditions developed previously sufficient for optimality. There are two main types of sufficiency theorems employed in optimal control theory. We refer to these loosely as the Mangasarian and the Arrow theorems. Actually, Arrow's original proof of his sufficiency theorem was incomplete although the theorem itself was correct. The correct proof of Arrow's sufficiency theorem is generally attributed to Seierstad and Sydsæter (1977).

Mangasarian's theorem essentially states that, when no state-space constraints are present, the Pontryagin necessary conditions are also sufficient if the Hamiltonian is convex (when minimizing) with respect to *both* the state and the control variables. By contrast, the Arrow sufficiency theorem requires only that the Hamiltonian expressed in terms of the optimal controls be convex with respect to the state variables.

3.3.8.1 The Mangasarian Theorem

We are interested in this section in proving one version of the Mangasarian (1966) sufficiency theorem for the continuous-time optimal control problem. This can be done with relative ease for the case of fixed initial and terminal times. We will additionally assume that there are no terminal time conditions and that the initial state is known and fixed. We will also assume that the Hamiltonian, when minimizing, is jointly convex in both the state variables and the control variables.

In particular, we study the following version of Mangasarian's theorem articulated by Seierstad and Sydsæter (1977) and Seierstad and Sydsæter (1999):

Theorem 3.7. *Restricted Mangasarian sufficiency theorem. Suppose the admissible pair* (x^*, u^*) *satisfies all of the relevant continuous-time optimal control problem necessary conditions for OCP*$(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ *when regularity in the sense of Definition 3.2 obtains, the set of feasible controls* U *is convex, the Hamiltonian* H *is jointly convex in* x *and* u *for all admissible solutions,* t_0 *and* t_f *are fixed,* x_0 *is fixed,* $K\left[x\left(t_f\right), t_f\right] = 0$*, and there are no terminal time conditions* $\Psi\left[x\left(t_f\right), t_f\right] = 0$*. Then any solution of the continuous-time optimal control necessary conditions is a global minimum.*

Proof. We follow the exposition of Seierstad and Sydsæter (1999) and begin the proof by noting that for (x^*, u^*) to be optimal it must be that

$$\Delta \equiv \int_{t_0}^{t_f} f_0\left(x, u, t\right) dt - \int_{t_0}^{t_f} f_0\left(x^*, u^*, t\right) dt \geq 0 \quad \forall \text{ admissible } (x, u) \tag{3.173}$$

when minimizing. Moreover, the associated Hamiltonian is

$$H = f_0 + \lambda^T \dot{x}$$

We note that (3.173) may be restated as

$$\Delta = \int_{t_0}^{t_f} \left(H - H^*\right) dt - \int_{t_0}^{t_f} \lambda^T \left(\dot{x} - \dot{x}^*\right) dt \tag{3.174}$$

Since H is convex with respect to x and u, the tangent line underestimates and we write

$$H^* + \frac{\partial H^*}{\partial x}\left(x - x^*\right) + \frac{\partial H^*}{\partial u}\left(u - u^*\right) \leq H$$

or equivalently

$$\frac{\partial H^*}{\partial x}\left(x - x^*\right) + \frac{\partial H^*}{\partial u}\left(u - u^*\right) \leq H - H^* \tag{3.175}$$

It follows from (3.174) and (3.175) that

$$\Delta \geq \int_{t_0}^{t_f} \left[\frac{\partial H^*}{\partial x}\left(x - x^*\right) + \frac{\partial H^*}{\partial u}\left(u - u^*\right)\right] dt - \int_{t_0}^{t_f} \lambda^T \left(\dot{x} - \dot{x}^*\right) dt \tag{3.176}$$

Using the adjoint equation $-d\lambda^T/dt = \partial H^*/\partial x$, this last result becomes

$$\Delta \geq \int_{t_0}^{t_f} \left[-\frac{d\lambda^T}{dt}\left(x - x^*\right) + \frac{\partial H^*}{\partial u}\left(u - u^*\right)\right] dt - \int_{t_0}^{t_f} \lambda^T \left(\dot{x} - \dot{x}^*\right) dt$$

$$= -\int_{t_0}^{t_f} \left[\frac{d\lambda^T}{dt} \left(x - x^* \right) + \lambda^T \left(\dot{x} - \dot{x}^* \right) \right] dt + \int_{t_0}^{t_f} \frac{\partial H^*}{\partial u} \left(u - u^* \right) dt$$

$$= -\int_{t_0}^{t_f} \frac{d}{dt} \left[\lambda^T \left(x - x^* \right) \right] dt + \int_{t_0}^{t_f} \frac{\partial H^*}{\partial u} \left(u - u^* \right) dt$$

$$= -\left[\lambda^T \left(x - x^* \right) \right]_{t_0}^{t_f} + \int_{t_0}^{t_f} \frac{\partial H^*}{\partial u} \left(u - u^* \right) dt \tag{3.177}$$

Expression (3.177) allows us to write

$$\Delta \geq -\left\{ \lambda^T \left(t_0 \right) \left[x \left(t_0 \right) - x^* \left(t_0 \right) \right] \right\} + \left\{ \lambda^T \left(t_f \right) \left[x \left(t_f \right) - x^* \left(t_f \right) \right] \right\}$$
$$+ \int_{t_0}^{t_f} \frac{\partial H^*}{\partial u} \left(u - u^* \right) dt$$

$$= -\lambda^T \left(t_0 \right) [0] + \left\{ 0 \left[x \left(t_f \right) - x^* \left(t_f \right) \right] \right\} + \int_{t_0}^{t_f} \frac{\partial H^*}{\partial u} \left(u - u^* \right) dt$$

$$= \int_{t_0}^{t_f} \frac{\partial H^*}{\partial u} \left(u - u^* \right) dt \geq 0 \tag{3.178}$$

where inequality (3.178) follows from the convexity of U and the fact the minimum principle is satisfied. Thus $\Delta \geq 0$, and we have established optimality. ■

Theorem 3.7 is easily extended to the case of mixed terminal conditions and a non-trivial salvage function. These generalizations are left as an exercise for the reader.

3.3.8.2 The Arrow Theorem

In Arrow and Kurz (1970) an alternative sufficiency theorem is presented, a theorem that is generally credited to Arrow. The Arrow theorem is based on a reduced form of the Hamiltonian obtained when the optimal control law derived from the minimum principle is employed to eliminate control variables from the Hamiltonian. Under appropriate conditions, if the reduced Hamiltonian is convex in the state variables, the necessary conditions are also sufficient. Seierstad and Sydsæter (1999) provide the following statement and proof of the Arrow result:

Theorem 3.8. *The Arrow sufficiency theorem. Let* (x^*, u^*) *be an admissible pair for OCP*$(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ *when regularity in the sense of Definition 3.2 obtains, the set of feasible controls is* $U = R^m$*, the Hamiltonian* H *is jointly convex in* x *and* u *for all admissible solutions,* t_0 *and* t_f *are fixed,* x_0 *is fixed,* $K\left[x\left(t_f\right), t_f\right] = 0$*, and there are no terminal time conditions* $\Psi\left[x\left(t_f\right), t_f\right] = 0$.

If there exists a continuous and piecewise continuously differentiable function $\lambda = (\lambda_1, \ldots, \lambda_n)^T$ *such that the following conditions are satisfied:*

$$\dot{\lambda}_i = \frac{-\partial H^*}{\partial x_i}, \quad \textit{almost everywhere} \qquad i = 1, \ldots, n \tag{3.179}$$

$$H(x^*, u, \lambda(t), t) \geq H(x^*, u^*, \lambda, t) \quad \textit{for all } u \in U \textit{ and all } t \in \left[t_0, t_f\right] \tag{3.180}$$

$$\widehat{H}(x, \lambda, t) = \min_{u \in U} H(x, u, \lambda, t) \textit{ exists and is convex in } x \textit{ for all } t \in \left[t_0, t_f\right] \tag{3.181}$$

then (x^*, u^*) *solves OCP*$(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ *for the given. If* $\widehat{H}(x, \lambda, t)$ *is strictly convex in* x *for all* t*, then* x^* *is unique (but* u^* *is not necessarily unique).*

Proof. Suppose (x, u) and (x^*, u^*) are admissible pairs and that (x^*, u^*) satisfies the minimum principle and related necessary conditions. Optimality will be assured if we show that

$$\Delta = \int_{t_0}^{t_f} f_0(x, u, t)\, dt - \int_{t_0}^{t_f} f_0\left(x^*, u^*, t\right) dt \geq 0 \quad \forall \text{ admissible } (x, u) \tag{3.182}$$

Suppose we are able to establish, for all admissible (x, u), that

$$H - H^* \geq \dot{\lambda}^T (x - x^*) \tag{3.183}$$

Then it is immediate that

$$\Delta = \int_{t_0}^{t_f} \left(H - H^*\right) dt - \int_{t_0}^{t_f} \lambda^T \left(\dot{x} - \dot{x}^*\right) dt \geq 0 \tag{3.184}$$

which in turn implies (3.182). Consequently, it is enough to establish condition (3.183). From definition (3.181) for $\widehat{H}$, we have

$$H^* = \widehat{H}^*$$
$$H \geq \widehat{H}$$

Therefore

$$H - H^* \geq \widehat{H} - \widehat{H}^* \tag{3.185}$$

Consequently it suffices to prove

$$\widehat{H} - \widehat{H}^* \geq -\dot{\lambda}^T (x - x^*) \tag{3.186}$$

for any admissible x, an inequality which makes $-\dot{\lambda}$ a subgradient of $\widehat{H}(x, \lambda, t)$ at x^*. To prove the existence of the subgradient, let us suppose

$$\widehat{H} - \widehat{H}^* \geq a^T (x - x^*), \tag{3.187}$$

again for all admissible x. If $\widehat{H}$ is differentiable, it is immediate that

$$\nabla \widehat{H_x}\Big|_{x=x^*} = \nabla H_x^* = -\dot{\lambda} \tag{3.188}$$

From (3.185) and (3.187), we have

$$H - H^* \geq a^T (x - x^*) \tag{3.189}$$

Consequently

$$G(x) = H - H^* - a^T (x - x^*) \geq 0 \quad \forall x \tag{3.190}$$

for any $t \in [t_0, t_f]$ and all admissible x. Note that since $G(x^*) = 0$, we know x^* minimizes $G(x)$. Therefore $\nabla_x G(x^*) = 0$; that is

$$\nabla g(x^*) = \nabla H|_{x=x^*} - a = 0, \tag{3.191}$$

Moreover, the adjoint equation compells

$$-\dot{\lambda} = \nabla H^* = a, \tag{3.192}$$

thereby establishing the existence of a subgradient, namely $a = \nabla H^*$. It is then immediate from (3.189) and (3.192) that

$$\widehat{H} - \widehat{H}^* \geq a^T (x - x^*) = -\dot{\lambda}(x - x^*) \tag{3.193}$$

which is recognized to be identical to (3.186) and completes the proof. ■

3.4 Optimal Control Examples

In this section, we provide simple examples of optimal control problems, solved using the necesary and sufficient conditions we have derived above.

3.4.1 Simple Example of the Minimum Principle

Let us employ the necessary and sufficient conditions developed previously to solve an illustrative continuous-time optimal control problem that has only upper and lower bound constraints on its controls. The example we select was originally proposed by Sethi and Thompson (2000):

$$\min J = \int_0^5 \left(\frac{1}{2} u^2 + 3u - x \right) dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= x + u \\ x\,(0) &= 10 \\ 0 &\le u \le 4 \end{aligned}$$

where x and u are scalars. Our first step is to form the Hamiltonian as the sum of the integrand plus an adjoint variable times the right-hand side of the state dynamics:

$$H = \frac{1}{2}u^2 + 3u - x + \lambda\,(x + u)$$

where λ is the adjoint variable for the state dynamics. The minimum principle requires

$$u = \left[\arg\left(\frac{\partial H}{\partial u} = 0\right)\right]_0^4$$

where the notation $[.]_a^b$ refers to the minimum norm projection operator for the interval $[a, b] \in \Re^1$ of the real line defined by

$$[v]_a^b = \begin{cases} b & \text{if } v \ge b \\ v & \text{if } a < v < b \\ a & \text{if } v \le a \end{cases}$$

Consequently

$$u = [\arg\,(u + 3 + \lambda = 0)]_0^4$$

or

$$u = [-3 - \lambda]_0^4$$

Furthermore, the adjoint dynamics are

$$(-1)\,\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = -1 + \lambda$$

with transversality condition

$$\lambda\,(5) = 0$$

Therefore we must solve the problem

$$\begin{aligned} \frac{d\lambda}{dt} &= 1 - \lambda \\ \lambda\,(5) &= 0 \end{aligned}$$

The exact solution is

$$\lambda\,(t) = 1 - \exp(5 - t)$$

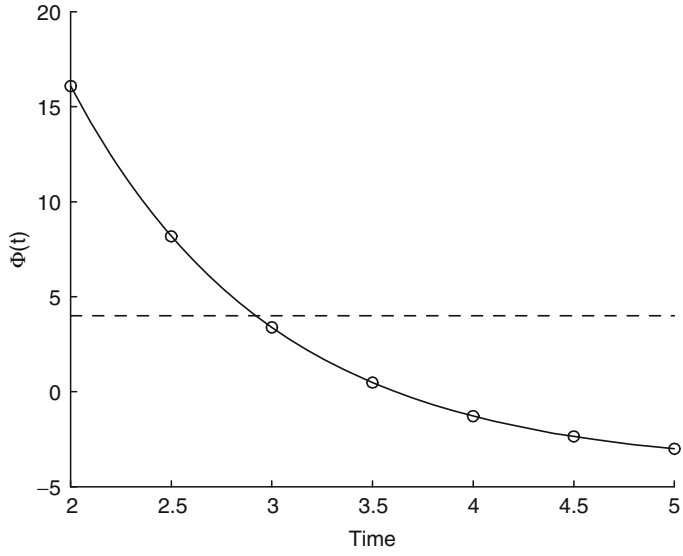


Fig. 3.2 Plot of the control law

with the consequence that the optimal control must obey

$$u = [\Phi]_0^4$$

where

$$\Phi \equiv -4 + e^{5-t}$$

The plot of the control law is given in Figure 3.2 where we see that $\Phi(t)$ exceeds the control upper bound once and crosses the time axis once for $t \in [0, 5]$. In fact, the instants in time for these phenomena are found by solving the following two equations:

$$\Phi(t) = 4 \Longrightarrow t = 5 - \ln 8 \approx 2.9206$$
$$\Phi(t) = 0 \Longrightarrow t = 5 - \ln 4 \approx 3.6137$$

The necessary conditions admit only one solution and are sufficient, due to joint convexity of the Hamiltonian in x and u; that solution is given by

$$u^* = \begin{cases} 4 & \text{for } t \in [0, 2.9206] \\ -4 + e^{5-t} & \text{for } t \in (2.9206, 3.6137] \\ 0 & \text{for } t \in (3.6137, 5] \end{cases}$$

3.4.2 An Example Involving Singular Controls

When optimal control problems are linear in their control variables, the necessary conditions may admit so-called bang-bang controls as well as singular controls. We will consider the following continuous-time optimal control problem, also originally proposed by Sethi and Thompson (2000):

$$\min J = \int_0^4 \frac{1}{2}x^2 dt$$

subject to

$$\begin{gathered} \frac{dx}{dt} = u \\ x\,(0) = 3 \\ -1 \leq u \leq +1 \end{gathered}$$

where x and u are scalars. We of course begin by forming the Hamiltonian:

$$H = \frac{1}{2}x^2 + \lambda u$$

The minimum principle requires that

$$u = \begin{cases} +1 & \text{for } \lambda < 0 \\ -1 & \text{for } \lambda > 0 \\ u_s & \text{for } \lambda\,(t) = 0 \end{cases}$$

In order for the singular control strategy u_s to be meaningful, the coefficient of the control u in the Hamiltonian, namely λ for this specific example, must vanish over a nontrivial arc of time $[t_1, t_2]$ where $t_2 > t_1$. Only then do we have a singular control. When the points in time at which the control coefficient vanishes are finite in number and countable, we may ignore them and develop a pure bang-bang control strategy wherein the optimal control is either at its upper bound or its lower bound. Note that even when there is no singular control the number of switchings between the upper and lower bounds cannot generally be known in advance. Neither may we know without some analysis whether the optimal control strategy begins at time t_0 with an upper bound or a lower bound. We use the name *control synthesis* to describe the process of determining the time intervals and order of bang-bang and singular controls.

For the present example the adjoint equation and boundary condition are

$$\begin{gathered} (-1)\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = x \\ \lambda\,(3) = 0 \end{gathered}$$

We see immediately that the adjoint dynamics are coupled to the state dynamics in that we cannot find the adjoint variable unless we know the state variable. Moreover, we cannot know the state variable unless we know the control, and we cannot know the control unless we know the adjoint variable. To break out of this simultaneity of the conditions describing the state, control and adjoint variables, we posit that

$$u = -1 \text{ for } t \in [0, t_1)$$

This is a wise choice since the criterion will be minimized when the state vanishes and the chosen control strategy reduces the state variable value from its initial value of unity at the maximal feasible rate. We realize that to reduce the state below zero would be inefficient due to the quadratic nature of the integrand of the criterion functional. The state initial-value problem for $t \in [0, t_1)$ is

$$\frac{dx}{dt} = -1 \quad x(0) = 3$$

Consequently

$$x = 3 - t$$

We further posit, based on the argument given above, that

$$x(t_1) = 3 - t_1 = 0 \quad \text{for} \quad t \in [0, t_1)$$

which requires that

$$t_1 = 3$$

Since the state will have reached its ideal value of zero at time t_1, we are inclined to believe that

$$u = u_s = 0 \text{ for } t \in [3, 4]$$

To check that our candidate solution

$$u^* = \begin{cases} -1 & \text{for } t \in [0, 3) \\ 0 & \text{for } t \in [3, 4] \end{cases} \tag{3.194}$$

satisfies the necessary conditions, we need to find the adjoint variable and show that it has appropriate signs/values on the two time intervals of interest.

In particular, we note that for $t \in [3, 4]$ the state dynamics are the initial-value problem

$$\frac{dx}{dt} = 0 \quad x(3) = 0 \tag{3.195}$$

with solution

$$x(t) = 0 \quad \text{for} \quad t \in [3, 4]$$

A direct consequence is that we have the following terminal-value problem for the adjoint variable:

$$\frac{d\lambda}{dt} = -x = 0 \quad \lambda(4) = 0 \quad \text{for } t \in [3, 4]$$

Therefore

$$\lambda(t) = 0 \quad \text{for } t \in [3, 4]$$

which is consistent with the singular control strategy (3.194).

Now we note that the adjoint dynamics and transversality condition for $t \in [0, 3)$ are

$$\frac{d\lambda}{dt} = -x = -3 + t \quad \lambda(3) = 0\,,$$

and therefore

$$\lambda(t) = -3t + \frac{1}{2}t^2 + \frac{9}{2} \quad \text{for } t \in [0, 3) \tag{3.196}$$

The graph of this adjoint variable as a function of time is given in Figure 3.3, which makes clear that $\lambda(t)$ given by (3.196) is positive for all $t \in [0, 3)$, as required for the posited solution (3.194) to satisfy the necessary conditions. Since the Hamiltonian is jointly convex in (x, u), (3.194) is the optimal control strategy for this example.

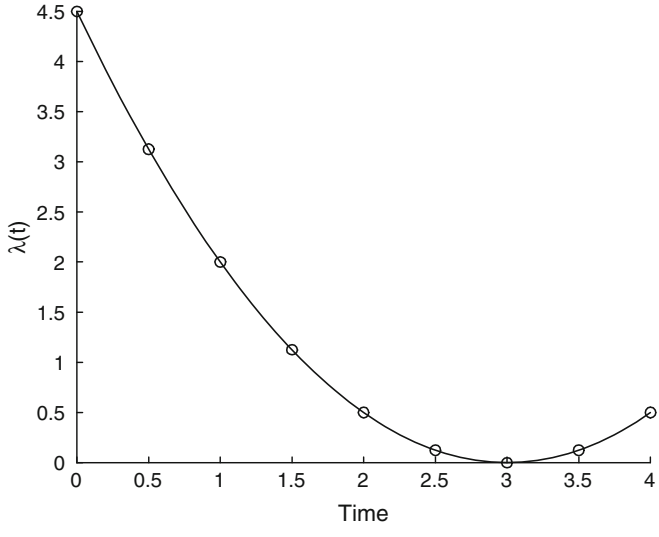


Fig. 3.3 Graph of adjoint variable as a function of time

3.4.3 Approximate Solution of Optimal Control Problems by Time Discretization

The preceeding example, namely

$$\min J = \int_0^4 \frac{1}{2} x^2 dt \tag{3.197}$$

subject to

$$\frac{dx}{dt} = u \tag{3.198}$$

$$x(0) = 3 \tag{3.199}$$

$$-1 \leq u \leq +1 \tag{3.200}$$

where x and u are scalars may be solved by making a discrete time approximation that may in turn be solved by finite-dimensional mathematical programming algorithms. In fact, if we employ N finite time intervals normalized to have identical unit length, we obtain the following mathematical program:

$$\min \; J = \sum_{t=0}^{N-1} \frac{1}{2} x_t^2$$

subject to

$$\begin{aligned} x_{t+1} &= x_t + u_t \qquad \forall t = 0, 1, \ldots, N-1 \\ x_0 &= 3 \\ -1 \leq u_t &\leq +1 \qquad \forall t = 0, 1, \ldots, N \end{aligned}$$

We are able to recognize the singular control of the optimal trajectory shown in (3.194) only when an appropriately fine level of time resolution is employed. That is, some levels of temporal resolution may not be able to find the singular control strategy that characterizes the optimal control for problem (3.197), (3.198), (3.199), and (3.200); such a circumstance is illustrated in Figure 3.4. In particular, we see that the jump discontiuity from $u^* = -1$ to $u^* = 0$ that occurs at $t = 3$ is disguised when N is small and becomes increasingly obvious as N grows.

3.4.4 A Two-Point Boundary-Value Problem

As we see from the preceding examples the minimum principle will frequently lead to a control law in the form of an equation that relates the optimal control strategy to the state and adjoint variables. This control law may sometimes be used to

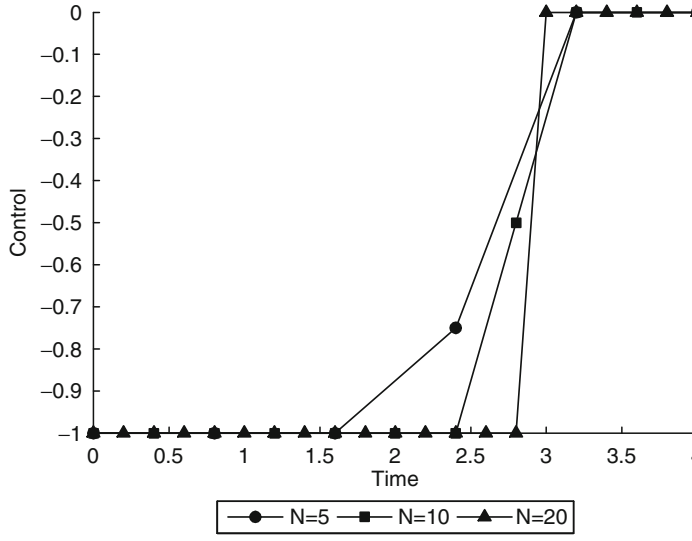


Fig. 3.4 Optimal control trajectories with different levels of time resolution

eliminate the control variables from both the state dynamics and the adjoint dynamics. When this is the case, we are left with a so-called two-point boundary-value problem of the form

$$\frac{dx}{dt} = F(x, \lambda, t) \tag{3.201}$$

$$\frac{d\lambda}{dt} = G(x, \lambda, t) \tag{3.202}$$

$$x(t_0) = x_0 \tag{3.203}$$

$$\lambda(t_f) = \lambda_f \tag{3.204}$$

where vector notation is employed. Furthermore t_0, t_f, x_0, and λ_f are known. This problem is called a two-point boundary-value problem because some of the variables sought have boundaries defined at time t_0 and the rest have boundaries defined at time t_f. A general method of solution whose convergence is assured does not exist for such problems.

Nonetheless, methods exist for the direct solution of (3.201), (3.202), (3.203), and (3.204) under appropriate regularity conditions. It is not our intent to present a formal treatment of numerical methods for two-point boundary-value problems, as we believe other methods to be presented in later chapters are easier to implement and offer distinct advantages. It is nonetheless important for any practitioner of optimal control theory to understand the key notions surrounding the numerical solution of two-point boundary-value problems. Accordingly, we now illustrate the class of

numerical methods known as *shooting methods* via a simple example. Consider the following problem:

$$\min J = \int_0^1 \frac{1}{2}\left(x^2 + u^2\right) dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= u \\ x\,(0) &= 1.5431 \end{aligned}$$

where x and u are scalars. We note that

$$\begin{aligned} H &= \frac{1}{2}\left(x^2 + u^2\right) + \lambda u \\ -\frac{d\lambda}{dt} &= \frac{\partial H}{\partial x} = x \\ \lambda\,(1) &= 0 \end{aligned}$$

Furthermore

$$u = \arg\left(\frac{\partial H}{\partial u} = 0\right) = \arg\left(u + \lambda = 0\right)$$

Therefore

$$u = -\lambda$$

and

$$\frac{dx}{dt} = -\lambda$$

Consequently the two-point boundary-value problem that expresses the necessary conditions is

$$\begin{aligned} \frac{d\lambda}{dt} &= -x \\ \frac{dx}{dt} &= -\lambda \\ x\,(0) &= 1.5431 \\ \lambda\,(1) &= 0 \end{aligned}$$

We can guess the initial condition for the adjoint variable; in fact let us assume

$$\lambda\,(0) = R^k$$

where R^k is a scalar denoting the initial value of the adjoint variable for iteration k of the shooting method. Thus, we will repeatedly solve the initial-value problem

$$\frac{d\lambda}{dt} = -x \tag{3.205}$$

$$\frac{dx}{dt} = -\lambda \tag{3.206}$$

$$x(0) = 1.5431 \tag{3.207}$$

$$\lambda(0) = R^k \tag{3.208}$$

We will be pleased if one of the "shots fired" by solving the single-point initial-value problem hits its metaphorical "target," namely that it satisfies

$$\lambda(1) \approx 0$$

If not, we adjust the value of $\lambda^k(0)$, then fire another shot. Table 3.1 shows the progression of a heuristic shooting algorithm wherein the missing adjoint initial condition is adjusted heuristically based on changes in the sign of $\lambda^k(1)$. The attentive reader may have noticed that a shooting method is not actually needed for the present example since (3.205), (3.206), (3.207), and (3.208) may be restated by noting that

$$\frac{d^2x}{dt^2} = -\frac{d\lambda}{dt} = x$$

and

$$\frac{dx(1)}{dt} = -\lambda(1) = 0$$

Hence, we could instead solve the following second-order differential equation with explicit boundary conditions:

$$\frac{d^2x}{dt^2} = x$$

$$x(0) = 1.5431$$

$$\frac{dx(1)}{dt} = 0$$

Table 3.1 Progression of heuristic shooting method

Iteration k	$\lambda^k(0) = R^k$	$\lambda^k(1)$
1	.5	-1.0419
2	1	-0.2704
3	1.5	$+0.5012$
4	1.25	$+0.1154$
5	1.125	-0.0775
⋮	⋮	⋮
10	1.1752	$-6.2058 \times 10^{-6} \approx 0$

whose exact solution is

$$x(t) = 0.18394e^{t} + 1.3592e^{-t}$$

It follows that

$$\begin{aligned} \frac{d\lambda}{dt} &= -0.18394e^{t} - 1.3592e^{-t} \\ \lambda(1) &= 0 \end{aligned}$$

for which the solution is

$$\lambda(t) = -0.18394e^{t} + 1.3592e^{-t}$$

Using the solution just obtained for the adjoint, it is an easy matter to verify the transversality condition:

$$\lambda(1) = -2.9096 \times 10^{-10} \approx 0$$

The above solution obtained from reduction of the optimality conditions to a single second-order differential equation is readily seen to agree with that found using the heuristic shooting algorithm. However, it is essential to note that reduction of the two-point boundary-value problem to a system of second-order differential equations with appropriate boundary conditions is neither generally nor typically possible.

3.4.5 Example with Free Terminal Time

Consider the following problem:

$$\min J = \int_0^{t_f} \frac{1}{2}\left(x^2 + u^2\right) dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= u \\ -1 \le u &\le +1 \\ x(0) &= 1.5431 \\ x(t_f) &= 1 \\ t_f &\ \text{free} \end{aligned}$$

where x and u are scalars. By introducing the dual variable θ, we price out the terminal constraint to obtain the alternative form

$$\min J_1 = \theta \cdot \left[x(t_f) - 1\right] + \int_0^{t_f} \frac{1}{2}\left(x^2 + u^2\right) dt$$

for which it is immediate that

$$\begin{aligned} H &= \frac{1}{2}\left(x^2 + u^2\right) + \lambda u \\ -\frac{d\lambda}{dt} &= \frac{\partial H}{\partial x} = x \\ \lambda\left(t_f\right) &= \frac{\partial \theta\left[x(t_f) - 1\right]}{\partial x(t_f)} = \theta \end{aligned}$$

Furthermore, without control constraints the minimum principle requires

$$\begin{aligned} u &= \arg\left(\frac{\partial H}{\partial u} = 0\right) = \arg\left(u + \lambda = 0\right) \\ &\Longrightarrow u = -\lambda \end{aligned}$$

However, owing to the upper and lower bound constraints, we must employ a projection, and so we write

$$u = [-\lambda]_{-1}^{+1}$$

Given the need to decrease the state variable $x(t)$ from 1.5431 to 1.0, it is reasonable to attempt a solution based on the control strategy

$$u = -1 \quad \forall t \in \left[0, t_f\right] \tag{3.209}$$

so that

$$\frac{dx}{dt} = -1 \Longrightarrow x = 1.5431 - t$$

Furthermore, because t_f is not fixed, we have

$$\left(H + \frac{\partial \Psi}{\partial t}\right)_{t=t_f} = 0 \Longrightarrow \frac{1}{2}\left(\left[x\left(t_f\right)\right]^2 + u^2\right) + \lambda\left(t_f\right) u = 0$$

However

$$\frac{1}{2}\left(\left[x\left(t_f\right)\right]^2 + u^2\right) + \lambda\left(t_f\right) u = \frac{1}{2}\left([1]^2 + [-1]^2\right) + \lambda\left(t_f\right)(-1) = 0$$

which requires

$$\lambda\left(t_f\right) = 1$$

That is, we now know

$$x\left(t_f\right) = \lambda\left(t_f\right) = \theta = 1$$

The two-point boundary-value problem that expresses the necessary conditions is

$$\begin{aligned} \frac{d\lambda}{dt} &= -x \\ \frac{dx}{dt} &= -\lambda \\ x\,(0) &= 1.5431 \\ \lambda\left(t_f\right) &= \theta = 1 \end{aligned}$$

It is obvious from the above that

$$\frac{d^2x}{dt^2} = -\frac{d\lambda}{dt} = x$$

and

$$\frac{dx(t_f)}{dt} = -\lambda\left(t_f\right) = -\theta$$

Hence, we may instead solve the second-order problem

$$\frac{d^2x}{dt^2} = x \tag{3.210}$$

$$x\,(0) = 1.543\,1 \tag{3.211}$$

$$\frac{dx(t_f)}{dt} = -\theta = -1 \tag{3.212}$$

Note that the solution must be of the form

$$x = Ae^t + Be^{-t}$$

Using this knowledge with (3.211), we have

$$A + B = 1.5431$$

By virtue of (3.212) we have

$$\frac{dx}{dt} = Ae^t - Be^{-t} \Longrightarrow Ae^{t_f} - Be^{-t_f} = -\theta = -1$$

Recalling that $x(t_f) = 1$, we also have

$$Ae^{t_f} + Be^{-t_f} = 1$$

Therefore, we must solve the system

$$\begin{aligned} A + B &= 1.5431 \\ Ae^{t_f} - Be^{-t_f} &= -1 \\ Ae^{t_f} + Be^{-t_f} &= 1 \end{aligned}$$

The solution is

$$\begin{aligned} A &= 0 \\ B &= 1.5431 \\ t_f &= 0.4338 \end{aligned}$$

That is

$$x = 1.5431e^{-t}$$

Note also that

$$\begin{aligned} \frac{d\lambda}{dt} &= -1.5431e^{t} \\ \lambda\,(0.4338) &= 1 \end{aligned}$$

whose exact solution is

$$\lambda\,(t) = 3.3812 - 1.5431e^{t}$$

and so

$$u = [-\lambda]_{-1}^{+1} = \left[1.5431e^{t} - 3.3812\right]_{-1}^{+1} = -1$$

as assumed in (3.209) since

$$\mathrm{F} = -\lambda = 1.54e^{t} - 3.38 \leq -1$$

for all $t \in [0, 0.434]$ as is made clear by Figure 3.5:

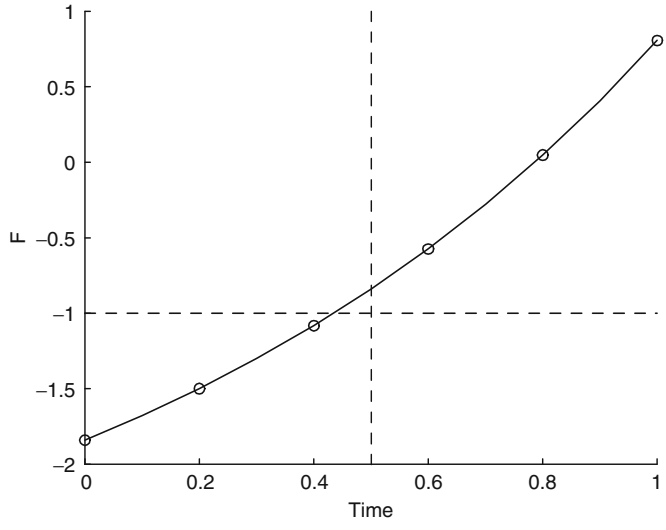


Fig. 3.5 Plot of $\mathrm{F} = (-1) \cdot \lambda$

3.5 The Linear-Quadratic Optimal Control Problem

Without a doubt one of the most significant versions of the continuous-time optimal control problem is that known as the *linear-quadratic problem* (LQP). This fame derives from the fact the Hamilton-Jacobi partial differential equation corresponding to the LQP can be solved very efficiently. The formulation we shall emphasize in this section corresponds to

$$\begin{aligned} f_0(x,u,t) &= \frac{1}{2}x^T A(t)\,x + \frac{1}{2}u^T B(t)\,u \\ f(x,u,t) &= F(t)\,x + G(t)\,u \\ K\left[x\left(t_f\right),t_f\right] &= \frac{1}{2}\left[x\left(t_f\right)\right]^T S\left(t_f\right)x\left(t_f\right) \\ U &= V, \text{ the entire vector space} \end{aligned}$$

so that the problem we face is

$$\begin{aligned} \min J[x(t),u(t)] &= K\left[x\left(t_f\right),t_f\right] + \int_{t_0}^{t_f} f_0[x(t),u(t),t]\,dt \\ &= \frac{1}{2}\left(x^T S x\right)_{t=t_f} + \frac{1}{2}\int_{t_0}^{t_f}\left(x^T Ax + u^T Bu\right)dt \end{aligned} \tag{3.213}$$

subject to

$$\frac{dx}{dt} = f(x(t),u(t),t) \tag{3.214}$$

$$= Fx + Gu \tag{3.215}$$

$$x(t_0) = x_0 \tag{3.216}$$

where $t_0 \in \Re_+^1$ is known and we have suppressed time dependencies of the matrices A, B, F, G, and S. There is no terminal time constraint, and both the initial time (t_0) and the terminal time (t_f) are fixed. Likewise x_0 is fixed. Furthermore, we assume that the *matrices A and S are positive semidefinite and the matrix B is positive definite*.

3.5.1 LQP Optimality Conditions

It is a relatively simple matter to derive an equivalent two-point boundary-value formulation of this problem since the necessary conditions are also sufficient due to the assumptions of positive definiteness just mentioned. In fact we find

$$H(x,u,\lambda) = \frac{1}{2}\left(x^T Ax + u^T Bu\right) + \lambda^T(Fx + Gu) \tag{3.217}$$

$$-\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = Ax + F^T\lambda \tag{3.218}$$

$$u = \arg\left\{\frac{\partial H}{\partial u} = Bu + G^T\lambda = 0\right\} \tag{3.219}$$

$$\Longrightarrow u = -B^{-1}G^T\lambda \tag{3.220}$$

It follows at once that

$$\begin{bmatrix}\dot{x}\\ \dot{\lambda}\end{bmatrix} = \begin{bmatrix}F & -GB^{-1}G^T\\ -A & -F^T\end{bmatrix}\begin{bmatrix}x\\ \lambda\end{bmatrix} \tag{3.221}$$

with

$$x(t_0) = x_0 \tag{3.222}$$

$$\lambda(t_f) = \left[\frac{\partial K}{\partial x}\right]_{t=t_f} = \left[\frac{\partial}{\partial x}\left(x^T S x\right)\right]_{t=t_f} \tag{3.223}$$

Clearly, (3.221), (3.222), and (3.223) constitute a linear two-point boundary-value problem. As such we could solve this system using a shooting method.

Instead we attempt a direct solution of the Hamilton-Jacobi partial differential equation:

$$H^*\left(x, \frac{\partial J^*}{\partial x}\right) + \frac{\partial J^*}{\partial t} = 0 \tag{3.224}$$

where J^* is the optimal-value function and H^* is obtained by evaluating the Hamiltonian along its optimal trajectory using the control law (obtained from the minimum principle) and the identity

$$\lambda^T = \frac{\partial J^*}{\partial x} \Longrightarrow \lambda = \left(\frac{\partial J^*}{\partial x}\right)^T \tag{3.225}$$

also valid along the optimal trajectory. Thus

$$\begin{aligned} H^*\left(x, \frac{\partial J^*}{\partial x}\right) &\equiv \min_{u\in U}\left[H(x,u,\lambda)\right]_{\lambda=\left(\frac{\partial J^*}{\partial x}\right)^T}\\ &= \left[H(x,u,\lambda)\right]_{u=-B^{-1}G^T\lambda,\,\lambda=\left(\frac{\partial J^*}{\partial x}\right)^T}\\ &= \left[\frac{1}{2}\left(x^TAx+u^TBu\right)+\lambda^T(Fx+Gu)\right]_{u=-B^{-1}G^T\lambda,\lambda=\left(\frac{\partial J^*}{\partial x}\right)^T} \end{aligned}$$

$$= \frac{1}{2}x^T Ax + \frac{1}{2}\left[B^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T\right]^T B\left[B^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T\right]$$
$$+ \left(\frac{\partial J^*}{\partial x}\right)Fx + \left(\frac{\partial J^*}{\partial x}\right)G\left[-B^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T\right] \tag{3.226}$$

Note that

$$\left[B^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T\right]^T B\left[B^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T\right]$$
$$= \left[\left(\frac{\partial J^*}{\partial x}\right)^T\right]^T \left(G^T\right)^T \left(B^{-1}\right)^T \left(BB^{-1}\right)\left[G^T\left(\frac{\partial J^*}{\partial x}\right)^T\right]$$
$$= \left(\frac{\partial J^*}{\partial x}\right)G\left(B^{-1}\right)^T G^T\left(\frac{\partial J^*}{\partial x}\right)^T = \left(\frac{\partial J^*}{\partial x}\right)GB^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T \tag{3.227}$$

since $GB^{-1}G^T$ is a symmetric matrix. Results (3.226) and (3.227) give

$$H^*\left(x, \frac{\partial J^*}{\partial x}\right) = \frac{1}{2}x^T Ax + \left(\frac{\partial J^*}{\partial x}\right)Fx - \frac{1}{2}\left(\frac{\partial J^*}{\partial x}\right)GB^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T \tag{3.228}$$

so that the Hamilton-Jacobi partial differential equation is

$$\frac{1}{2}x^T Ax + \left(\frac{\partial J^*}{\partial x}\right)Fx - \frac{1}{2}\left(\frac{\partial J^*}{\partial x}\right)GB^{-1}G^T\left(\frac{\partial J^*}{\partial x}\right)^T + \frac{\partial J^*}{\partial t} = 0 \tag{3.229}$$

with

$$J\left[x\left(t_f\right), t_f\right] = \frac{1}{2}\left[x\left(t_f\right)\right]^T S\left(t_f\right)x\left(t_f\right) \tag{3.230}$$

as the boundary condition.

3.5.2 The HJPDE and Separation of Variables for the LQP

We attempt a solution of (3.229) subject to (3.230) by employing the following transformation:

$$J^* = \frac{1}{2}x^T Z\left(t\right)x \tag{3.231}$$

where $Z\left(t\right)$ is an unknown time-dependent symmetric matrix. (This is a very specific instance of a technique known as *separation of variables* commonly used to

solve so-called *separable partial differential equations*.) Substituting (3.231) into (3.229) yields after some manipulation

$$\frac{1}{2}x^T\left[\frac{dZ}{dt}+ZF+F^TZ-ZGB^{-1}G^TZ+A\right]x=0 \tag{3.232}$$

Hence

$$\frac{dZ}{dt}+ZF+F^TZ-ZGB^{-1}G^TZ+A=0 \tag{3.233}$$

which is known as the *matrix Ricatti equation* and is subject to the boundary condition

$$Z\left(t_f\right)=S\left(t_f\right) \tag{3.234}$$

in keeping with the original boundary condition (3.230).

3.5.3 LQP Numerical Example

Consider the example familiar from our discussion of the shooting method in Section 3.4.4, namely

$$\min J=\int_0^1\frac{1}{2}\left(x^2+u^2\right)dt \tag{3.235}$$

subject to

$$\frac{dx}{dt}=u \tag{3.236}$$

$$x\left(0\right)=1.5431 \tag{3.237}$$

where x and u are scalars. In terms of the notation of the previous section, we have

$$\begin{aligned}A&=1\\B&=1\\F&=0\\G&=1\\K&=0\\S&=0\\t_0&=0\\t_f&=1\end{aligned}$$

This means that the Ricatti equation

$$\frac{dZ}{dt}+ZF+F^TZ-ZGB^{-1}G^TZ+A=0$$

becomes

$$\frac{dZ}{dt} - Z^2 + 1 = 0 \tag{3.238}$$

$$Z(1) = 0 \tag{3.239}$$

with solution given by

$$Z(t) = -\tanh(t-1) \tag{3.240}$$

so that

$$Z(0) = .76159 \tag{3.241}$$

We find the initial adjoint variable by noting

$$\begin{aligned}\lambda(0) &= \left[\frac{\partial J^*}{\partial x}\right]_{t=0} = \left\{\frac{\partial}{\partial x}\left[\frac{1}{2}x(t)^T Z(t) x(t)\right]\right\}_{t=0} \\ &= Z(0) x(0) = .76159(1.5431) = 1.1752\end{aligned} \tag{3.242}$$

which is immediately recognized as the value that allowed the shooting method to converge when applied to this problem in Section 3.4.4.

3.5.4 Another LQP Example

Consider the problem

$$\min J = \int_0^1 \frac{1}{2}\left(x^2 + u^2\right) dt \tag{3.243}$$

subject to the following dynamics that, while still linear, involve both state and control variables on the righthand side:

$$\frac{dx}{dt} = x + u \quad (\lambda) \tag{3.244}$$

$$x(0) = 1 \tag{3.245}$$

We know that

$$\begin{aligned}H &= \frac{1}{2}\left(x^2 + u^2\right) + \lambda(x + u) \\ -\frac{d\lambda}{dt} &= \frac{\partial H}{\partial x} = x + \lambda \\ \lambda(1) &= 0 \\ u &= \arg\left[\frac{\partial}{\partial u} H = u + \lambda = 0\right] \\ &\Longrightarrow u = -\lambda\end{aligned}$$

Note that this is a linear-quadratic problem for which

$$\frac{dZ}{dt} + ZF + F'Z - ZGB^{-1}G'Z + A = 0 \tag{3.246}$$

$$Z\left(t_f\right) = S\left(t_f\right) \tag{3.247}$$

$$A = B = F = G = 1 \tag{3.248}$$

$$S = 0 \tag{3.249}$$

$$t_f = 1 \tag{3.250}$$

Consequently

$$\frac{dZ}{dt} + 2ZF - Z^2 + 1 = 0$$

$$Z\left(1\right) = 0$$

whose exact solution is

$$Z\left(t\right) = \frac{1}{2}\left(\sqrt{2} - 2\tanh\left(t - \frac{1}{4}\sqrt{2}\left(\ln\frac{\sqrt{2}-1}{\sqrt{2}+1} + 2\sqrt{2}\right)\right)\sqrt{2}\right)\sqrt{2} \tag{3.251}$$

which in turn tells us that

$$Z\left(0\right) = 1.6895 \tag{3.252}$$

It is then immediate that

$$\lambda\left(0\right) = Z\left(0\right)x\left(0\right) = 1.6895\left(1\right) = 1.6895$$

Thus, we have the following initial-value problem:

$$\begin{aligned}
\frac{dx}{dt} &= x + u = x - \lambda \\
-\frac{d\lambda}{dt} &= \frac{\partial H}{\partial x} = x + \lambda \\
x\left(0\right) &= 1 \\
\lambda\left(0\right) &= 1.6895
\end{aligned}$$

Numerical solution of the above initial-value problem gives

$$\begin{pmatrix} \lambda\left(0\right) \\ \lambda\left(.25\right) \\ \lambda\left(.5\right) \\ \lambda\left(.75\right) \\ \lambda\left(1\right) \end{pmatrix} = \begin{pmatrix} 1.6895 \\ 1.1097 \\ .67012 \\ .31516 \\ 1.3022 \times 10^{-6} \end{pmatrix} \tag{3.253}$$

Note that $\lambda(1) \approx 0$. Since $u = -\lambda$, we have the following optimal control solution:

$$\begin{pmatrix} u(0) \\ u(.25) \\ u(.5) \\ u(.75) \\ u(1) \end{pmatrix} = \begin{pmatrix} -1.6895 \\ -1.1097 \\ -.670\,2 \\ -.315\,6 \\ \approx 0 \end{pmatrix} \tag{3.254}$$

3.6 Exercises

1. Prove Theorem 3.5.
2. Give a formal statement of a theorem that establishes the Euler-Lagrange equations derived in Section 3.1.4 are, in fact, valid necessary conditions for the problem given by (3.26), (3.27), and (3.28). Be sure to include all regularity conditions in your theorem.
3. Give a formal statement and proof of Intiligator's duality theorem, mentioned in Section 3.1.8, for isoperimetric constraints in the calculus of variations.
4. Numerically solve the brachistochrone problem of Section 3.2.3 using $N = 20$, 50, 100. Comment on your findings.
5. Extend the Mangasarian sufficiency theorem of Section 3.3.8.1 to include consideration of terminal constraints

$$\Psi\left[x\left(t_f\right), t_f\right] = 0 \in \Re^r$$

and nontrivial terminal costs $K\left[x\left(t_f\right), t_f\right]$.
6. Construct and prove an optimal control sufficiency theorem that includes consideration of mixed state and control constraints.
7. Describe how an optimal control problem with free terminal time may be approached and solved using a sequence of problems with fixed terminal time. Apply your method to the example problem of Section 3.4.5.
8. Consider the problem

$$\min J = \frac{1}{2}[x(1)]^2 + \int_0^1 \left(\frac{1}{2}x^2 + u\right) dt \tag{3.255}$$

subject to

$$\frac{dx}{dt} = \frac{1}{8}x - u \quad (\lambda) \tag{3.256}$$

$$-1 \leq u \leq +1 \tag{3.257}$$

$$x(0) = 1 \tag{3.258}$$

where both x and u are scalars. Do the Arrow and/or Mangasarian sufficiency theorems hold? Solve by the minimum principle.

9. Create an optimal control problem that has two control variables, a singular control that is optimal, and satisfies the Arrow sufficiency theorem. Establish that the singular control is optimal by applying the necessary conditions for continuous-time optimal control.
10. Consider the problem

$$\min J = \frac{1}{2}[x(1)]^2 + \int_0^1 \frac{1}{2}\left(x^2 + u^2\right) dt \tag{3.259}$$

subject to

$$\frac{dx}{dt} = x + u \quad (\lambda) \tag{3.260}$$

$$x(0) = 1 \tag{3.261}$$

where both x and u are scalars. Solve by a shooting method.
11. Solve the problem of Exercise 9 above by forming and solving the Hamilton-Jacobi partial differential equation.

List of References Cited and Additional Reading

Arrow, K. J. and M. Kurz (1970). *Public Investment, the Rate of Return ,and Optimal Fiscal Policy*. Baltimore: The Johns Hopkins University Press.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton: Princeton University Press.

Bellman, R. E. (1961). *Adaptive Control Processes: A Guided Tour*. Princeton: Princeton University Press.

Bernoulli, J. (1697a). Curvatura radii in diaphanis nonuniformibus [the curvature of a ray in nonuniform media]. *Acta Eruditorum*, 206–211.

Bernoulli, J. (1697b). Solutio problematum fraternorum. una cum propositione reciproca aliorum [solution of problems of my brother. together with the proposition of others in turn]. *Acta Eruditorum*, 211–217.

Bliss, G. (1946). *Lectures on the Calculus of Variations*. Chicago: University of Chicago Press.

Brechtken-Manderscheid, U. (1991). *Introduction to the Calculus of Variations*. London: Chapman & Hall.

Chiang, A. C. (1992). *Elements of Dynamic Optimization*. McGraw-Hill.

Dreyfus, S. E. (1965). *Dynamic Programming and the Calculus of Variations*. New York: Academic Press.

Gelfand, I. M. and S. V. Fomin (2000). *Calculus of Variations*. Mineola, NY: Dover.

Hestenes, M. R. (1966). *Calculus of Variations and Optimal Control Theory*. New York: John Wiley.

Intriligator, M. D. (1971). *Mathematical Optimization and Economic Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Mangasarian, O. L. (1966). Sufficient conditions for the optimal control of nonlinear systems. *SIAM Journal of Control 4*, 19–52.

Miele, A. (1962). The calculus of variations in applied aerodynamics and flight mechanics. In G. Leitmann (Ed.), *Optimization Techniques*. New York: Academic Press.

Seierstad, A. and K. Sydsæter (1977). Sufficient conditions in optimal control theory. *International Economic Review 18*, 367–391.

Seierstad, A. and K. Sydsæter (1999). *Optimal Control Theory with Economic Applications*. Amsterdam: Elsevier.

Sethi, S. and G. Thompson (2000). *Optimal Control Theory: Applications to Management Science and Economics*. MA: Kluwer Academic Publishers.

Spruck, J. (2006). *Course Notes on Calculus of Variations*. Baltimore: The Johns Hopkins University.

Tait, K. (1965). *Singular Problems in Optimal Control*. Ph. D. thesis, Harvard University, Cambridge, MA.

Thornton, S. T. and J. B. Marion (2003). *Classical Dynamics of Particles and Systems*. Pacific Grove, CA: Brooks Cole.

Valentine, F. A. (1937). The problem of lagrange with differential inequalities as added side conditions. In *Contributions to the Theory of the Calculus of Variations, 1933–1937*. Chicago: University of Chicago Press.

Venkataraman, P. (2001). *Applied Optimization with MATLAB Programming*. New York: Wiley-Interscience.

Chapter 4
Infinite Dimensional Mathematical Programming

In this chapter we are concerned with the generalization of finite-dimensional mathematical programming to infinite-dimensional vector spaces. This topic is pertinent to dynamic optimization because dynamic optimization in continuous time *de facto* occurs in infinite-dimensional spaces since the variable $x(t)$, even if x is a scalar, has an infinity of values for continuous $t \in [t_0, t_f] \subseteq \Re^1_+$ where $t_f > t_0$.

In fact, in this chapter we will define a class of dynamic optimization problems that is more general than the calculus of variations and optimal control problems we have discussed up to this point. We will then show how the foundation material of this chapter allows us to derive the Euler-Lagrange equation and the Pontryagin minimum principle from the notions of a Gâteaux derivative and variational inequality optimality conditions for infinite-dimensional mathematical programs. In this way we are able, despite the introductory nature of this book, to understand the deep connection between infinite-dimensional mathematical programing and continuous-time optimal control without resort to discrete-time approximations.

We also derive Kuhn-Tucker type necessary conditions for infinite-dimensional mathematical programs in preparation for the extension of mathematical programming algorithms familiar from the study of finite-dimensional nonlinear programming to Hilbert spaces. In fact, because of the connection between infinite-dimensional mathematical programming and continuous-time optimal control, we will be able to apply the algorithms developed in this chapter to a variety of models discussed in subsequent chapters.

The following is an outline of the contents of this chapter:

Section 4.1: Elements of Functional Analysis. We present some elementary properties of topological vector spaces, emphasizing the differences between finite-dimensional and infinite-dimensional vector spaces.

Section 4.2: Variational Inequalities and Constrained Optimization of Functionals. We present the notion of an infinite dimensional mathematical program and show that it has necessary conditions in the form of infinite dimensional variational inequalities.

Section 4.3: Continuous-Time Optimal Control. We derive necessary conditions for optimal control problems from necessary conditions for infinite dimensional mathematical programs.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_4, © Springer Science+Business Media, LLC 2010

Section 4.4: Optimal Control with Time Shifts. We extend the results of the previous section to problems involving time shifts.

Section 4.5: Derivation of the Euler-Lagrange Equation. We derive the Euler-Lagrange equation from necessary conditions for infinite dimensional mathematical programs.

Section 4.6: Kuhn-Tucker Conditions for Hilbert Spaces. In this section, we show that, under appropriate regularity conditions, Kuhn-Tucker conditions may be articulated for infinite dimensional mathematical programs.

Section 4.7: Mathematical Programming Algorithms. Having established optimality conditions for infinite dimensional mathematical programs, we set about expressing and testing continuous-time algorithms that are direct generalizations of algorithms familiar from nonlinear programming in finite-dimensional spaces.

4.1 Elements of Functional Analysis

This chapter contains several important results from functional analysis that are stated without proof in order to focus the reader's energies on those aspects of infinite-dimensional mathematical programming that are essential to model building and solution without lengthy detours. At the end of this chapter we provide a list of references that may be consulted should the reader wish to study the formal proofs omitted here.

4.1.1 Notation and Elementary Concepts

To proceed the reader will need to recall some basic notions from elementary analysis, including the following which are defined without elaboration:

1. **Norm**: The norm of a vector v shall be denoted as $\|v\|$. The norm itself is a pre-established notion of "distance" or "length."
2. **Neighborhood**: Given a point $x \in S$, a set of interest, and $\varepsilon \in \Re^1_{++}$, the set

$$N_\varepsilon(x) = \{y : ||y - x|| \leq \varepsilon\}$$

   is called the *ε-ball* or *ε-neighborhood* of $x \in S$.
3. **Bounded Set**: A set S is *bounded* if it can be enclosed in a ball of finite radius.
4. **Closed Set**: Given a set S, the *closure* of S, denoted cl S, is the set of points that are arbitrarily close to S; that is $x \in$ cl S if, for each $\varepsilon > 0$,

$$S \cap N_\varepsilon(x) \neq \emptyset$$

   where $N_\varepsilon(x) = \{y : ||y - x|| \leq \varepsilon\}$. The set S is *closed* if $S = cl\ S$.

5. **Open Set**: A set that is not closed is an *open* set.
6. **Compact Set**: A set S is *compact* if it is closed and bounded.
7. **Fundamental Set Operations**: Given two sets A and B, with $A \subset S$ and $B \subset S$, the following are fundamental set operations:

$$\begin{aligned} A \cap B &= \{x : x \in A \text{ and } x \in B\} \\ A \cup B &= \{x : x \in A \text{ or } x \in B\} \\ A \setminus B &= \{x : x \in A \text{ and } x \notin B\} \\ A + B &= \{z = x + y : x \in A \text{ and } y \in B\} \\ A - B &= \{z = x - y : x \in A \text{ and } y \in B\} \end{aligned}$$

   Note that $A \subset S$ implies that if $x \in A$, then $x \in S$. In our exposition, this will be taken to mean A is a proper ($A \neq S$) subset of S.
8. **Interior Point**: Given $S \subset V$, a normed vector space, then the point $a \in S$ is an *interior point* of S if there is an $\epsilon > 0$ such that all vectors x satisfying $\|x - a\| < \epsilon$ are also members of S. The set of all interior points of S is denoted *int* S.
9. $O(n)$ **Notation**: Given two sequences $\{a_n\}$ and $\{b_n\}$ such that $b_n \geq 0$ for all n, we say $\{a_n\}$ is of order $\{b_n\}$ and write

$$a_n = O\,[b_n] \text{ as } n \to \infty$$

   when

$$\lim_{n \to \infty} \frac{a_n}{b_n} = 0$$

4.1.2 Topological Vector Spaces

We need to define some basic concepts and introduce some key results from functional analysis, which is the branch of mathematics that deals with problems whose solutions are functions of a continuous independent variable; that continuous independent variable is usually time. As remarked previously, because any nontrivial subinterval of continuous time – sometimes referred to as an "arc of time" – contains an uncountable number of instants of time, we say continuous-time problems are infinite-dimensional. There are some subtleties associated with infinite-dimensional analysis, and certainly the notion of convergence is the most important of these. In particular the notion of pointwise convergence and the notion of convergence in the sense of operators are not equivalent and lead to distinct topologies. For this reason, infinite-dimensional vector spaces are sometimes called *topological vector spaces*. Moreover, as we shall see, fundamental constructions – like the gradient – have different realizations in different topological vector spaces. For these reasons, analyses of continuous-time problems require the clear and unambiguous articulation of the vector spaces in which one is working. Note also that we will ultimately

be concerned – when we return to direct consideration of the calculus of variations and optimal control – solely with normed spaces for which the norm is induced by an inner product. Thus, in applications, when we speak of a specific normed topological vector space, the reader will need to recall the unique, unambiguous, well-defined inner product associated with it.

In our discussion of the foundations of topological vector spaces that follows, we assume the reader already has some familiarity with what is meant by a *normed vector space* on $\Re^n$; consequently our explanations of many of the most basic concepts are given in a narrative style and are neither detailed nor illustrated by examples. Furthermore, we repeat a point made previously: some of the theoretical results for infinite-dimensional vector spaces contained in the ensuing overview are presented without proof so that nonmathematicians may penetrate quickly to the level appropriate for their applications. The reader with prior exposure to functional analysis and topological vector spaces may wish to skip to Section 4.2 of this chapter.

Loosely speaking, a *vector space* V is a nonempty set of elements for which vector addition and multiplication of vectors by scalars are defined. A *normed vector space* on $\Re$ is a vector space V for which a mapping $V \longrightarrow \Re^1$ called the norm is defined. We reiterate that the *norm* of $v \in V$ may be thought of as the "length" of v, and that it is denoted by $\|v\|$. The relevant formal definition is:

Definition 4.1. *Normed linear vector space. A normed linear vector space V is a vector space V on which there is a well-defined real-valued function that maps each $v \in V$ into a real number $\|v\|$ called the norm of v. The norm has the following properties:*

1. $\|v\| \geq 0 \quad \forall v \in V$ *with* $\|v\| = 0 \Longleftrightarrow v = 0$
2. $\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$
3. $\|\lambda v\| = |\lambda| \, \|v\| \quad \forall v \in V, \lambda \in \Re^1$.

Note that the above three properties may be referred to as the zero length property, the triangle inequality property, and the scalar multiplication property, respectively.

One of the most important vector spaces in applied mathematics is that of continuous functions. Another is the space of functions of bounded variation. To provide examples and also to prepare for subsequent analyses, we now formally define these spaces and some associated concepts:

Definition 4.2. *Space of continuous functions. The normed linear vector space $C[a, b]$ consists of continuous functions on an interval of the real line $[a, b]$ where $a, b \in \Re^1$ and $a < b$. The norm of $x \in C[a, b]$ is*

$$\|x\| = \max \{|x(t)| : a \leq t \leq b\}$$

Definition 4.3. *Bounded variation. For an interval of the real line $[a, b]$ where $a, b \in \Re^1$ and $a < b$, the partition created by the finite set of points $t_i \in [a, b]$ for $i = 0, 1, 2, \ldots, n$ such that*

$$a = t_0 < t_1 < t_2, \ldots < t_n = b$$

gives rise to the bounded variation of a function $f(x)$ defined on $[a, b]$ provided there exists a finite constant K such that

$$\sum_{i=1}^{n} [x(t_i) - x(t_{i-1})] \leq K$$

Definition 4.4. *Total variation. The total variation of a function x defined on the real interval $[a, b]$ is denoted by*

$$TV(x) = \int_a^b |dx(t)|$$

and defined by

$$TV(x) = \sup \sum_{i=1}^{n} [x(t_i) - x(t_{i-1})]$$

for a given partition of $[a, b]$.

Definition 4.5. *Space of functions of bounded variation. The normed linear vector space $BV[a, b]$ consists of all functions of bounded variation on an interval of the real line $[a, b]$ where $a, b \in \Re^1$ and $a < b$. Its norm is*

$$\|x\| = |x(a)| + TV(x)$$

Note that the total variation of a constant is zero.

The vector spaces in which we are mainly interested, as the preceding examples suggest, are spaces of functions. In the definitions and results that follow, the vector spaces employed should be considered spaces of functions unless there is an explicit statement to the contrary. In this and subsequent chapters, we will have cause to speak of *mappings* – sometime called *transformations* or *operators*. Also among the definitions introduced below is the definition of a functional. We have already introduced these concepts in an informal way in previous chapters; now we give the following formal definitions:

Definition 4.6. *Mapping. Let V and W be vector spaces and let D be a subspace of V. A rule that associates an element $y \in \mathcal{W}$ with every element of $x \in D$ is said to be a mapping F from V to W with domain D, and we write*

$$F : D \longrightarrow W$$

or $y = F(x)$.

We will have cause to refer to the related notion of a linear form defined on a vector space:

Definition 4.7. *Linear form. We say a linear and continuous mapping from a vector space V to $\Re^1$ is a continuous linear form.*

From prior discussions, we already are familiar with the concept of a functional:

Definition 4.8. *Functional. A mapping F from a vector space V into $\Re^1$, the space of real numbers, is called a functional, and we write*

$$F : V \longrightarrow \Re^1$$

We may allow functionals to be vectors that arise when mapping from V to $\Re^n$, so that the following definition obtains:

Definition 4.9. *Vector functional. A mapping J from a vector space V into $\Re^n$, the space of finite-dimensional n-vectors of real numbers, is called a vector functional, and we write*

$$J : V \longrightarrow \Re^n$$

Note that the norm of an infinite-dimensional vector space is a functional. Furthermore, both scalar and vector functionals are by definition mappings and often referred to as transformations or operators.

Next we take care to note that there is a notion of *strong convergence* that we distinguish from *weak convergence*:

Definition 4.10. *Strong convergence. In a normed vector space V, we shall say that the sequence $\left\{v^k \in V\right\}$ converges strongly to $v \in V$ as $k \longrightarrow \infty$ if and only if*

$$\lim_{k \longrightarrow \infty} \left\| v^k - v \right\| = 0$$

Definition 4.11. *Weak convergence. Let V be a normed vector space and V^* its dual. We say that the sequence $\left\{v^k \in V\right\}$ converges weakly to $v \in V$ as $k \longrightarrow \infty$ if and only if*

$$\lim_{k \longrightarrow \infty} \mathcal{F}\left(v^k\right) = \mathcal{F}(v)$$

for all $\mathcal{F} \in V^*$.

The dual space referred to in Definition 4.11 is explained subsequently in Definition 4.24. Furthermore, the dichotomy of strong convergence and weak convergence, introduced in Definitions 4.10 and 4.11, is recognized by the terminology set forth in the following definitions of the two fundamental topologies:

Definition 4.12. *Strong topology. A vector space for which the property of strong convergence holds throughout is said to exhibit the strong topology.*

Definition 4.13. *Weak topology. A vector space for which the property of weak but not strong convergence holds throughout is said to exhibit the weak topology.*

It is not hard to prove the following key result:

Theorem 4.1. *Relationship of strong and weak convergence. If $v^k \longrightarrow v$ strongly, then $v^k \longrightarrow v$ weakly.*

It is fundamental that the converse of this theorem is true in finite-dimensional spaces (specifically, in $\Re^n$) but is generally FALSE in infinite-dimensional spaces. So the two topologies, weak and strong, are distinct. This leads to strong and weak topological distinctions for every mathematical concept that relies on the notion of convergence of sequences.

The two notions of convergence (topology) also naturally give rise to two notions of continuity:

Definition 4.14. *Strong continuity. We say that the functional J is strongly continuous if*

$$\left[v^k \longrightarrow v \text{ strongly}\right] \Longrightarrow J\left(v^k\right) \longrightarrow J\left(v\right)$$

Definition 4.15. *Weak continuity. Similarly we say that the functional J is weakly continuous if*

$$\left[v^k \longrightarrow v \text{ weakly}\right] \Longrightarrow J\left(v^k\right) \longrightarrow J\left(v\right)$$

We introduce next the notion of semicontinuity:

Definition 4.16. *Strong lower semicontinuity. We say that the functional $J : V \longrightarrow \Re^1$ is strongly lower semicontinuous if, for all $v \in V$ and for all sequences $\left\{v^k\right\} \in V$, the implication of $v^k \rightarrow v$ (strongly) is that*

$$\liminf J\left(v^k\right) \geq J\left(v\right) \tag{4.1}$$

Definition 4.17. *Weak lower semicontinuity. We say that the functional $J : V \longrightarrow \Re^1$ is weakly lower semicontinuous if, for all $v \in V$ and for all sequences $\left\{v^k\right\} \in V$, the implication of $v^k \rightarrow v$ (weakly) is that*

$$\liminf J\left(v^k\right) \geq J\left(v\right) \tag{4.2}$$

Similarly, strong and weak upper semicontinuity are defined by replacing (4.1) and (4.2) by

$$\limsup J\left(v^k\right) \leq J\left(v\right)$$

We will also make use of the notion of uniform continuity:

Definition 4.18. *Uniform continuity. When $u, v \in V$, we say the functional J is uniformly continuous on V if for every $\varepsilon > 0$ there exists $\eta\left(\varepsilon\right) > 0$ such that*

$$\begin{aligned} \left\|u - v\right\| &\leq \eta\left(\varepsilon\right) \\ &\Longrightarrow \left|\left\langle J\left(u\right) - J\left(v\right), \phi\right\rangle\right| \leq \varepsilon \quad \forall \phi \text{ such that } \left\|\phi\right\| = 1 \end{aligned}$$

while

$$\eta\left(\varepsilon\right) \longrightarrow 0_+ \Longleftrightarrow \varepsilon \longrightarrow 0_+$$

We also define the notion of Lipschitz continuity:

Definition 4.19. *Lipschitz continuity. We say that the functional J is Lipschitz continous on V if there exists a positive constant $L \in \Re^1_{++}$ such that*

$$\|J(u) - J(v)\| \leq L \|u - v\|$$

for all $u, v \in V$.

We also have

Definition 4.20. *Cauchy sequence. A Cauchy sequence in V relative to the strong topology is a sequence $\{v^k \in V\}$ such that for every $\varepsilon > 0$ there exists $k(\varepsilon)$ such that*

$$\left\| v^l - v^m \right\| < \varepsilon \quad \forall l, m \geq k(\varepsilon)$$

Definition 4.21. *Complete vector space. The space V is said to be complete if every Cauchy sequence in V has a limit which is an element of V.*

Completeness is highly desirable as it assures that convergent algorithms yield meaningful results, in the sense of those results being within the space for which the underlying problem of interest is defined. Some topological vector spaces that are intuitively appealing are not complete.

Subsequently, we will need Lebesgue's dominated convergence theorem, which we now state:

Theorem 4.2. *Lebesgue's dominated convergence theorem. Let $\{f_n(u(t))\}$ be a sequence of measurable functions on V such that*

$$f(u(t)) = \lim_{n \longrightarrow \infty} f_n(u(t))$$

exists for every $u(t) \in V$. Suppose there exists an integrable function g such that

$$|f_n(u(t))| \leq g(u(t)) \quad n = 1, 2, 3, \ldots \quad \forall u \in V$$

Then

$$\lim_{n \longrightarrow \infty} \int_{t_0}^{t_f} f_n(u(t))\, dt = \int_{t_0}^{t_f} f(u(t))\, dt$$

Proof. See Rudin (1987). ■

We now offer the following definition of a specific category of vector spaces, namely Banach spaces, that play a critical role in the study of infinite-dimensional optimization problems:

Definition 4.22. *Banach space. A Banach space is a normed and complete vector space for the strong topology.*

We shall be concerned principally with Banach spaces; frequently we will be concerned with a special type of Banach space called a Hilbert space. Moreover, the mappings between vector spaces we shall consider in applications will generally be linear and frequently strongly continuous in the sense of the following definition:

Definition 4.23. *Linear mapping. If V and Y are two normed vector spaces on $\Re^1$, then the mapping $A : V \longrightarrow Y$ is called a linear mapping if for all $w, v \in V$ and for all $\lambda, \mu \in \Re^1$*

$$A(\lambda w + \mu v) = \lambda A(w) + \mu A(v)$$

Note that in an infinite-dimensional space a linear mapping is not necessarily continuous, unlike finite-dimensional spaces.

We denote the set of all linear and strongly continuous mappings from $V \longrightarrow Y$ by $\mathcal{L}(V, Y)$. Note that $\mathcal{L}(V, Y)$ is a vector space. It is not difficult to prove that a linear mapping $A : V \longrightarrow Y$ is strongly continuous if and only if there exists a constant $M \in (0, \infty)$ such that

$$\|A(v)\|_Y < M \|v\|_V$$

in the strong topology. In light of this property, it is possible to associate with every element of $\mathcal{L}(V, Y)$ (that is, with every linear, strongly continuous mapping $A : V \longrightarrow Y$) the real number

$$\|A\|_{\mathcal{L}(V,Y)} = \sup_{v \in V, v \neq 0} \frac{\|A(v)\|_Y}{\|v\|_V} \tag{4.3}$$

It can be formally shown that the operator $\|.\|_{\mathcal{L}(V,Y)}$ is a norm for the vector space $\mathcal{L}(V, Y)$. It may also be proven that if V is a normed vector space and Y a Banach space, then $\mathcal{L}(V, Y)$ is a Banach space.

With the above material concerning the space of all linear and strongly continuous mappings as preamble, we make the following definition:

Definition 4.24. *Dual of a normed vector space. If V is a normed vector space, the strong topological dual of V is $\mathcal{L}\left(V, \Re^1\right)$, the space of all linear and strongly continuous mappings from V to $\Re^1$.*

It is essential to note that the dual of a topological vector space is a space of functionals; in particular, every element of the dual space is a linear functional. It is immediate from Definition 4.24 that the (strong topological) dual of V, which we call V^*, is a Banach space. Furthermore, for $\mathcal{F} \in V^*$ the value of $\mathcal{F}$ at $v \in V$ is denoted by $\mathcal{F}(v)$, and the norm of $\mathcal{F}$ in V^* is defined as

$$\|\mathcal{F}\|_{V^*} = \sup_{v \in V, v \neq 0} \frac{|\mathcal{F}(v)|}{\|v\|_V}$$

We let V^{**} denote the dual of the dual V^*; furthermore, V^{**} is defined, is also a Banach space, and is called the *bidual of* V (in the sense of the strong topology).

Furthermore, we say that a Banach space is reflexive if $V^{**} = V$. There is an important result on reflexive Banach spaces called the weak compactness theorem, whose statement is the following:

Theorem 4.3. *Weak compactness theorem. If V is a reflexive Banach space, then from every bounded set of elements of V, it is possible to extract a subsequence converging weakly to an element of V.*

Proof. See Wouk (1979). ■

A very important theorem regarding Banach spaces is the contraction mapping theorem:

Theorem 4.4. *Contraction mapping theorem. Let V be a Banach space, Λ a metric space, and let $\Phi : \Lambda \times V \longrightarrow V$ be a Lipschitz continuous mapping with Lipschitz constant $L < 1$. Then, for each $\lambda \in \Lambda$ there exists a unique fixed point $v(\lambda) \in V$ such that*

$$v(\lambda) \in \Phi(\lambda, v(\lambda))$$

Moreover, the map $\lambda \longrightarrow v(\lambda)$ is continuous, and, for any $\lambda \in \Lambda$ and $u \in V$, we have

$$\|u - v(\lambda)\| \leq \frac{1}{1-L} \|u - \Phi(\lambda, u)\|$$

Proof. See Bressan and Piccoli (2007). ■

We next define the notion of a Hilbert space:

Definition 4.25. *Hilbert space. A vector space V with a scalar product $\langle ., . \rangle$ is called a Hilbert space if V is complete for the strong topology and norm $\|v\| = [\langle v, v \rangle]^{1/2}$.*

As a consequence, a Hilbert space is a Banach space for which the norm derives from the notion of a scalar product. Furthermore, every Hilbert space is reflexive. Another key result for Hilbert spaces is the following version of the representation theorem due to Riesz:

Theorem 4.5. *Riesz representation theorem. Let V be a Hilbert space and let $\mathcal{F} \in V^*$ be a continuous linear form on V. Then there exists a unique element $w^0 \in V$ such that*

$$\mathcal{F}(v) = \langle w^0, v \rangle \quad \forall v \in V$$

and

$$\|\mathcal{F}\|_{V^*} = \|w^0\|_V$$

Conversely, it is possible to associate with each $u \in V$ the continuous linear form $\mathcal{F}_u$ defined by

$$\mathcal{F}_u(v) = \langle u, v \rangle \quad \forall v \in V$$

There is an alternative form of the Riesz representation theorem given by Luenberger (1969) that is useful in some applications:

Theorem 4.6. *Alternative form of the Riesz representation theorem. Let f be a bounded linear functional on the space of continuous functions $C[a,b]$. Then there is a function v of bounded variation on $[a,b]$ such that for all $x \in X$*

$$f(x) = \int_a^b x(t)\, dv(t) \tag{4.4}$$

Moreover, the norm of f is the total variation of v on $[a,b]$. Furthermore, a bounded linear functional on $C[a,b]$ is defined by (4.4) for every function v of bounded variation on $[a,b]$.

4.1.3 Convexity

We begin our discussion of convexity in functional analysis with the following definition familiar from the study of finite-dimensional spaces:

Definition 4.26. *Convex set. A set S is convex if, for each $v^1, v^2 \in S$, and $\lambda \in [0,1]$, we have*

$$\lambda v^1 + (1-\lambda) v^2 \in S$$

Related notions are the convex hull and the closed convex hull:

Definition 4.27. *Convex hull. The convex hull of a set $S \subset V$, is the intersection of all convex sets in V containing S, and is denoted by $\{S\}$. The convex hull is the smallest convex set containing S. The set S is convex if and only if $S = \{S\}$.*

Definition 4.28. *Closed convex hull. The closed convex hull of a set $S \subset V$, is the intersection of all closed convex sets in V containing S, and is denoted by $cl\,\{S\}$. The set S is closed and convex if and only if $S = cl\,\{S\}$.*

These definitions set the stage for presentation and proof of the following theorem, portions of which are familiar from Chapter 2:

Theorem 4.7. *Properties of convexity. The following are true:(a) S is convex $\Rightarrow$ $cl\ S$ is convex; (b) S is convex $\Rightarrow$ $\{S\}$ is convex; (c) $\{S\}$ is convex $\Rightarrow$ $cl\ \{S\}$ is convex; (d) for convex sets A and B, the set $A \cap B$ is convex; and (e) for convex sets A and B, the set $A + B$ is convex.*

Proof. (a) Define the open ball $N_\varepsilon(v) = \{w : ||w - v|| < \varepsilon\}$ and two arbitrary points $v, w \in S$, a convex set. Let $u = \lambda v + (1-\lambda) w$, $0 \le \lambda \le 1$, and assume $v, w \in cl\ S$. Therefore, there exist points $v^0 \in N_\varepsilon(v) \cap S$ and $w^0 \in N_\varepsilon(w) \cap S$, and it is desired to prove that

$$u^0 = \lambda v^0 + (1-\lambda) w^0 \in cl\ S \Rightarrow u^0 \in N_\varepsilon(u)$$

Note that

$$\begin{aligned}\left\|u-u^0\right\| &= \left\|\lambda v+(1-\lambda)\,w-\left(\lambda v^0+(1-\lambda)\,w^0\right)\right\| \\ &= \left\|\lambda\left(v-v^0\right)+(1-\lambda)\left(w-w^0\right)\right\| \\ &\leq \lambda\left\|v-v^0\right\|+(1-\lambda)\left\|w-w^0\right\|\end{aligned}$$

Note $v^0 \in N_\varepsilon(v) \Longrightarrow \left\|v-v^0\right\| < \epsilon$ and $w^0 \in N_\varepsilon(w) \Longrightarrow \left\|w-w^0\right\| < \epsilon$; thus

$$\left\|u-u^0\right\| < \lambda\varepsilon+(1-\lambda)\epsilon = \epsilon$$

It is clear that $u^0 \in N_\varepsilon(u)$. Since u^0 is on the line segment joining v^0 and w^0 and S is convex by definition, then $u^0 \in S$. Therefore, $N_\varepsilon(u) \cap S \neq \emptyset$ for all $u \in S$ and $cl\ S$ is convex.

(b) $\{S\}$ is convex by definition.
(c) $cl\ \{S\}$ is convex by (a).
(d) If the S_i for $i = [1,n]$ are convex, then $\cap_j S_j \subset S_i$ for any $i \in [1,n]$. Since a contiguous subset of a convex set is clearly convex, $\cap_j S_j$ is convex.
(e) By the definition of convexity, $\lambda v^1+(1-\lambda)\,v^2 \in S$ for all $v^1, v^2 \in S$ where $\lambda \in (0,1)$, which implies closure under addition and scalar multiplication. ■

In some situations it is necessary to recognize and exploit the notion of local convexity, and hence we make the following definition:

Definition 4.29. *Convexity of linear spaces. A linear vector space is locally convex if it has a basis about* 0 *consisting entirely of convex sets.*

Note that a Banach space is locally convex.

4.1.4 The Hahn-Banach Theorem

The ability to separate sets using a linear functional is established by an important theorem in functional analysis known as the Hahn-Banach theorem, which is crucial to the development of optimality conditions. To articulate the Hahn-Banach theorem we must introduce some additional formal definitions. In particular:

Definition 4.30. *$\mathcal{H}$ is a hyperplane if there exists a nonzero continuous linear functional J and a real constant α such that*

$$\mathcal{H} = \{v \in V : J(v) = \alpha\}$$

Definition 4.31. *The hyperplane $\mathcal{H} = \{v \in V : J(v) = \alpha\}$, where $J : V \to V^*$ is a continuous linear mapping, is a separating hyperplane of A and B if one of the following holds:*

$$\begin{aligned}(i)\ J(a) &\leq \alpha \quad \forall a \in A \text{ and } J(b) \geq \alpha \quad \forall b \in B \\ (ii)\ J(a) &\geq \alpha \quad \forall a \in A \text{ and } J(b) \leq \alpha \quad \forall b \in B\end{aligned}$$

Definition 4.32. *The hyperplane $\mathcal{H} = \{v \in V : J(v) = \alpha\}$, where $J : V \to V^*$is a continuous linear mapping, is a strictly separating hyperplane of A and B if one of the following holds:*

$$(i)\ J(a) < \alpha \quad \forall a \in A \text{ and } J(b) > \alpha \quad \forall b \in B$$
$$(ii)\ J(a) > \alpha \quad \forall a \in A \text{ and } J(b) < \alpha \quad \forall b \in B$$

Definition 4.33. *Disjoint sets. We shall say two sets are disjoint if they share no interior points.*

The so-called geometrical version of the Hahn-Banach theorem may now be stated as follows:

Theorem 4.8. *Hahn-Banach theorem. Let A and B be disjoint nonempty convex sets in a vector space V. If $int\ A \neq \emptyset$, then there exists a hyperplane separating A and B.*

There are two further theorems important to the characterization of optimality that are directly related to the Hahn-Banach theorem, namely:

Theorem 4.9. *Weak separation theorem. Let A and B be disjoint nonempty convex sets in a vector space V. If A is open, then there exists a closed hyperplane separating A and B such that*

$$J(a) \geq \alpha \geq J(b) \quad \forall a \in A, \forall b \in B$$

where $J : V \to V^*$is a continuous linear mapping.

Theorem 4.10. *Strong separation theorem. Let A and B be disjoint nonempty convex sets in a locally convex vector space V. If A is compact, then given real numbers α and $\epsilon > 0$, there exists a closed hyperplane strictly separating A and B such that*

$$J(a) \geq \alpha > \alpha - \epsilon \geq J(b) \quad \forall a \in A, \forall b \in B$$

where $J : V \to V^*$is a continuous linear mapping. Furthermore $A \cap B = \emptyset$, $A \cap \mathcal{H} = \emptyset$, and $B \cap \mathcal{H} = \emptyset$.

4.1.5 Gâteaux Derivatives and the Gradient of a Functional

The derivative of a function $f(x)$ of a real variable $x \in \Re^1$ is of course defined by

$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

provided the indicated limit exists. This definition is not readily generalized to form the derivative of a functional $J(v)$ in a Banach space, since the decision variables v

may have kinked trajectories or exhibit jump discontinuities, which make the functional $J(v)$ nonsmooth. Consequently, in order to articulate necessary conditions for optimality in infinite-dimensional mathematical programming we rely on the notion of a directional derivative. In fact, in this section and in Section 4.1.6, we introduce two notions of a directional derivative in Banach space.

We first consider the Gâteaux derivative, or G-derivative, which is defined as follows:

Definition 4.34. *G-derivative. Let V be a vector space and J a functional defined on V. Provided it exists, the limit*

$$\lim_{\theta \longrightarrow 0 \in \Re^1} \frac{J(v+\theta\phi) - J(v)}{\theta} \equiv \delta J(v,\phi) \tag{4.5}$$

is the Gâteaux derivative (G-derivative) at $v \in V$ in the direction $\phi \in V$. If the limit $\delta J(v,\phi)$ exists for all $\phi \in V$, we say that the functional J is differentiable in the sense of Gâteaux (G-differentiable) at $v \in V$.

The G-derivative is a generalization of the idea of a directional derivative familiar in finite-dimensional spaces. The notion of a G-derivative allows us to give a very general characterization of the gradient of a functional in Banach spaces with a well-defined inner product (i.e., in a Hilbert space). In fact the following definition obtains:

Definition 4.35. *Gradient in Hilbert space. Let V be a Hilbert space with associated scalar product $\langle .,. \rangle$ and J a functional that is G-differentiable on V. The element $\nabla J(v) \in V$ such that*

$$\delta J(v,\phi) = \langle \nabla J(v), \phi \rangle \quad \forall \phi \in V$$

is called the gradient of J at v.

Note that the Riesz representation theorem for Hilbert spaces, Theorem 4.5, ensures that $\nabla J(v)$ exists and is well defined.

The preceding definition of the gradient of a functional is so important that we need to give an illustrative example:

Example 4.1. Distinct derivatives in distinct Hilbert spaces. Consider $x = (x_1, x_2, \ldots, x_n)^T$ along with mappings $v(x)$ belonging to the Sobolev space

$$\mathcal{H}^1[a,b] = \left\{ v : v \in L^2[a,b], \frac{\partial v}{\partial x_i} \in L^2[a,b],\ i = 1,2,\ldots,n \right\}$$

whose scalar product for $w, v \in \mathcal{H}^1[a,b]$ is

$$\langle w, v \rangle = \int_a^b \left[w^T \cdot v + \left(\frac{dw}{dt} \right)^T \cdot \frac{dv}{dt} \right] dt \tag{4.6}$$

where $[a,b] \in \Re^1$ is a closed interval of the real line and $L^2[a,b]$ is the space of square integrable functions for the same interval. It can be shown that $\mathcal{H}^1[a,b]$ is a Hilbert space. For the space of square integrable functions $L^2[a,b]$, the scalar (inner) product of $w, v \in L^2[a,b]$ is

$$\langle w, v \rangle = \int_a^b w^T \cdot v dt \tag{4.7}$$

Note that $L^2[a,b]$ is also a Hilbert space. Our interest in this example is finding a continuous linear form which is the gradient of the functional

$$J(v) = \int_a^b v^2 dt \tag{4.8}$$

for v in both $\mathcal{H}^1[a,b]$ and $L^2[a,b]$. To this end, we take the limit

$$\begin{aligned}
\delta J(v,\phi) &= \lim_{\theta \longrightarrow 0 \in \Re^1} \frac{J(v+\theta\phi) - J(v)}{\theta} \\
&= \lim_{\theta \longrightarrow 0 \in \Re^1} \int_a^b \frac{(v+\theta\phi)^2 - v^2}{\theta} dt \\
&= \lim_{\theta \longrightarrow 0 \in \Re^1} \int_a^b \frac{v^2 + 2\theta v^T \cdot \phi + (\theta\phi)^2 - v^2}{\theta} dt \\
&= \lim_{\theta \longrightarrow 0 \in \Re^1} \int_a^b \left[\frac{2\theta v^T \cdot \phi}{\theta} + \frac{\theta^2 \phi^2}{\theta} \right] dv \\
&= \int_a^b 2v^T \cdot \phi dv + \lim_{\theta \longrightarrow 0 \in \Re^1} \int_a^b \theta \phi^2 dv \\
&= \int_a^b 2v^T \cdot \phi dv
\end{aligned}$$

It is immediate from the identity $\delta J(v,\phi) = \langle \nabla J(v), \phi \rangle$ and the definition of the scalar product in $L^2[a,b]$ that

$$\begin{aligned}
\delta J(v,\phi) &= \int_a^b 2v^T \cdot \phi dv = \langle 2v, \phi \rangle \\
&\Longrightarrow \nabla J(v) = 2v
\end{aligned}$$

Moreover, in $\mathcal{H}^1[a,b]$, we can demonstrate that

$$\delta J(v,\phi) = \int_a^b 2v^T \cdot \phi dv \neq \langle 2v, \phi \rangle \tag{4.9}$$

In particular, in $\mathcal{H}^1[a,b]$, we have that

$$\langle 2v, \phi \rangle = \int_a^b \left[2v^T \cdot \phi + 2\left(\frac{dv}{dt}\right)^T \cdot \frac{d\phi}{dt} \right] dt$$

That is, in $\mathcal{H}^1[a,b]$, the G-derivative is not a continuous linear form consistent with $2v$ being the gradient of $J(v)$ defined by (4.8). More generally, the gradient is not invariant from one topological vector space to another.

4.1.6 The Fréchet Derivative

Another notion of directional derivative which is equivalent to the G-derivative in most cases is the Fréchet or F-derivative. The formal definition of the F-derivative is

Definition 4.36. *F-Derivative. Let V be a Banach space with dual V^* and J be a functional defined on V. The Fréchet derivative (F-derivative) of J at $v \in V$ in direction $\phi \in V$ is the continuous linear mapping $\Delta J(v,\phi) : V \times V \longrightarrow V^*$ such that*

$$\lim_{||\phi|| \to 0} \frac{||J(v+\phi) - J(v) - \Delta J(v,\phi)||}{||\phi||} = 0$$

If the limit $\Delta J\ (v,\phi)$ exists for all $\phi \in V$, we say that the functional J is differentiable in the sense of Fréchet (F-differentiable) at $v \in V$.

The following result establishes a relationship between the G-derivative and the F-derivative under very mild restrictions:

Theorem 4.11. *Relationship of Fréchet and Gâteaux derivatives. Given the real Banach space V and its dual V^*, as well as J a functional defined on V, if the Fréchet derivative $\Delta J(v,\phi) \in V^*$ is defined at $v \in V$ for direction ϕ, then the Gâteaux derivative $\delta J(v,\phi)$ is also defined at $v \in V$ for direction ϕ, and the two derivatives are equal.*

Proof. Since the Fréchet derivative is defined at V, we have

$$\lim_{||\phi|| \to 0} \frac{||J(v+\phi) - J(v) - \Delta J(v,\phi)||}{||\phi||} = 0$$

For any fixed nonzero $\phi \in V$, substitution of $\theta\phi$ for ϕ, where $\theta \to 0$ gives

$$\lim_{\theta \to 0} \frac{||J(v+\theta\phi) - J(v) - \Delta J(v,\theta\phi)||}{\theta\, \|\phi\|} = \lim_{\theta \to 0} \frac{||J(v+\theta\phi) - J(v) - \Delta J(v,\theta\phi)||}{\theta} = 0$$

Since $\Delta J(v, \theta\phi)$ is a continuous linear operator, $\Delta J(v, \theta\phi) = \theta\ \Delta J(v, \phi)$, and

$$\lim_{\theta \to 0} \frac{||J(v + \theta\phi) - J(v) - \Delta J(v, \theta\phi)||}{\theta} = \lim_{\theta \to 0} \frac{J(v + \theta\phi) - J(v)}{\theta} - \Delta J(v, \theta) = 0$$

or

$$\Delta J(v, \theta) = \lim_{\theta \to 0} \frac{J(v + \theta\phi) - J(v)}{\theta} = \delta J(v, \theta). \qquad \blacksquare$$

4.2 Variational Inequalities and Constrained Optimization of Functionals

We are ready to begin formal consideration of infinite-dimensional mathematical programs. For that consideration, we assume that we are given a topological vector space V and a functional $J : V \longrightarrow \Re^1$. We want to minimize J either on V or on some subset $U \subset V$. This fundamental problem includes finite-dimensional mathematical programming, as well as the classical calculus of variations and modern optimal control theory, as special cases.

To explore the constrained optimization of a functional, we recall one of the most famous of all theorems in mathematical analysis, namely the Weierstrass theorem:

Theorem 4.12. *Weierstrass existence theorem. If the subset $U \subset V$ is strongly (respectively weakly) compact and J is strongly (respectively weakly) continuous on U, then the problem*

$$\min J\ (v) \quad \text{s.t. } v \in U \subset V$$

has an optimal solution $v^* \in U$.

It is because of this theorem that optimization problems, both finite- and infinite-dimensional (static and dynamic), have intrinsic meaning and warrant systematic study. In infinite dimensions, the weak version of the Weierstrass theorem is used almost exclusively because of the difficulty associated with checking U for strong compactness. When the space V of interest is a reflexive Banach space or a Hilbert space, it is enough to establish that U is bounded and weakly closed to assure weak compactness.

We next establish that the G-derivative may be employed to state necessary conditions for an optimal solution. In particular, the following theorem obtains:

Theorem 4.13. *First-order necessary condition for unconstrained optimum. If $J\ (v)$ is a functional on V and is G-differentiable at $v^* \in V$, then a necessary condition for v^* to be an optimum of J is*

$$\delta J\ \left(v^*, \phi\right) = 0 \quad \forall \phi \in V$$

Proof. For $v^*$to be either a local or a global minimum of the functional J, it must be that the function $J(v^* + \theta\phi)$ of the real variable θ obeys

$$0 = \left[\frac{d}{d\theta} J\left(v^* + \theta\phi\right)\right]_{\theta=0} = \lim_{\theta \longrightarrow 0} \frac{J\left(v^* + \theta\phi\right) - J\left(v^*\right)}{\theta} \quad \forall \phi \in V$$

By (4.5) this limit is the G-derivative. ■

Another key result is:

Theorem 4.14. *Second-order necessary condition for an unconstrained minimum. If $J(v)$ is a functional and twice continuously differentiable at $v^* \in V$, necessary conditions for v^* to be a minimum of J are that*

$$\delta J\left(v^*, \phi\right) = 0 \quad \forall \phi \in V$$
$$\delta^2 J\left(v^*, \phi, \phi\right) \equiv \delta\left[\delta J\left(v^*, \phi\right), \phi\right] \geq 0 \quad \forall \phi \in V$$

Proof. This result follows directly from the second-order necessary condition for functions of a real variable θ. ■

We now turn our attention to the constrained minimization of functionals. The key result is:

Theorem 4.15. *Variational inequality necessary condition. Take $J(v)$ to be a functional on V that is G-differentiable at $v^* \in U$, and let $U \subset V$ be a convex set. A necessary condition for $v^* \in U$ to be a minimum of J on U is*

$$\delta J\left(v^*, v - v^*\right) \geq 0 \quad \forall v \in U \tag{4.10}$$

Proof. Consider an arbitrary $v \in V$. Since U is convex it must be that

$$\theta v + (1 - \theta) v^* = v^* + \theta\left(v - v^*\right) \in U \quad \forall \theta \in [0, 1]$$

For v^* to be a minimum of $J(v)$ subject to $v \in V$, it must be that $\theta = 0$ is a solution of the one-dimensional mathematical program

$$\min F(\theta) \equiv J\left[v^* + \theta\left(v - v^*\right)\right]$$

subject to

$$g_1(\theta) = -\theta \leq 0$$
$$g_2(\theta) = \theta - 1 \leq 0$$

for which the Kuhn-Tucker conditions are

$$\frac{dF(\theta)}{d\theta} + \zeta_1 \frac{dg_1(\theta)}{d\theta} + \zeta_2 \frac{dg_2(\theta)}{d\theta} = 0 \tag{4.11}$$
$$\zeta_1 g_1(\theta) = 0 \tag{4.12}$$

$$\zeta_2 g_2(\theta) = 0 \tag{4.13}$$

$$\zeta_1, \zeta_2 \geq 0 \tag{4.14}$$

At $\theta = 0$, we have $g_1 = 0$ and $g_2 < 0$, so that $\zeta_1 \geq 0$ and $\zeta_2 = 0$. It follows at once from the Kuhn-Tucker identity (4.11) that

$$\left[\frac{d}{d\theta} J\left(v^* + \theta\left(v - v^*\right)\right)\right]_{\theta=0} = \zeta_1 \geq 0 \tag{4.15}$$

This last result together with the definition of the G-derivative gives the variational inequality

$$\left[\frac{d}{d\theta} J\left(v^* + \theta\left(v - v^*\right)\right)\right]_{\theta=0} = \delta J\left(v^*, v - v^*\right) \geq 0 \quad \forall v \in V$$

since $v \in V$ was arbitrary in our development of (4.15). ■

Note that when the gradient of a functional J exists and U is a convex subset of a Hilbert space, the necessary condition (4.10) becomes the more familiar form

$$v^* \in U \quad \text{such that} \quad \left\langle \nabla J\left(v^*\right), v - v^* \right\rangle \geq 0 \quad \forall v \in U \tag{4.16}$$

This result follows from the representation theorem and the fact that $v - v^*$ for all $v \in U$ generates all direction vectors pointing away from v^* when U is convex.

4.3 Continuous-Time Optimal Control

In this section, we want to use the theory of infinite-dimensional mathematical programming developed above to give an alternative derivation of the necessary conditions for the continuous-time optimal control problem. We consider the following restricted form of the continuous-time optimal control problem:

$$\min J(u) = \int_{t_0}^{t_f} f_0(x, u, t)\, dt$$

subject to

$$\frac{dx}{dt} = f(x, u, t)$$

$$x(t_0) = \xi_0 \text{ (fixed)}$$

Note that there are no explicit control constraints. There are n state variables and m control variables; t_0, t_f, and $x(t_0)$ are fixed; there are neither terminal-time costs nor terminal-time constraints.

4.3.1 Analysis Based on the G-Derivative

We will assume

$$u \in \left(L^2\left[t_0, t_f\right]\right)^m \tag{4.17}$$

$$x\left(u,t\right) \;:\; \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \tag{4.18}$$

$$f_0 \;:\; \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow L^2\left[t_0, t_f\right] \tag{4.19}$$

$$f \;:\; \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n \tag{4.20}$$

Presently, we will assume the following properties of the state operator $x\left(u,t\right)$: it exists for all admissible u and is unique, strongly continuous and G-differentiable with respect to u. This means that the G-derivative

$$\delta x\left(u,\phi;t\right) = \lim_{\theta \longrightarrow 0} \frac{x\left(u+\theta\phi,t\right) - x\left(u,t\right)}{\theta} \equiv y\left(u\right)$$

exists. To simplify notation, we will sometimes use the symbol $y\left(u\right)$ to denote this G-derivative.

Note that $x\left(u,t\right)$ satisfies the integral equation

$$x\left(u,t\right) = x\left(t_0\right) + \int_{t_0}^{t} f\left[x\left(u,\xi\right), u, \xi\right] d\xi$$

from which it is immediate that

$$x\left(u+\theta\phi,t\right) = x\left(t_0\right) + \int_{t_0}^{t} f\left[x\left(u+\theta\phi,\xi\right), u+\theta\phi, \xi\right] d\xi$$

We note that the ratio

$$R(u,\phi;\theta) \equiv \frac{x\left(u+\theta\phi,t\right) - x\left(u,t\right)}{\theta}$$

has a numerator equivalent to

$$x\left(t_0\right) + \int_{t_0}^{t_f} f\left[x\left(u+\theta\phi,\xi\right), u+\theta\phi, \xi\right] d\xi - x\left(t_0\right) - \int_{t_0}^{t_f} f\left[x\left(u,\xi\right), u, \xi\right] d\xi$$

which may be simplified to obtain

$$R(u,\phi;\theta) = \int_{t_0}^{t} \frac{f\left[x\left(u+\theta\phi,\xi\right), u+\theta\phi, \xi\right] - f\left[x\left(u,\xi\right), u, \xi\right]}{\theta} d\xi$$

The last expression yields, upon taking the limit $\theta \longrightarrow 0$, the following:

$$\delta x\left(u,\phi;t\right)=\lim_{\theta\longrightarrow\mathbf{0}}R(u,\phi;\theta)=\int_{t_0}^{t}\delta f\left(u,\phi;\xi\right)d\xi$$

where

$$\delta f\left(u,\phi;\xi\right)\equiv\lim_{\theta\longrightarrow\mathbf{0}}\frac{f\left[x\left(u+\theta\phi,\xi\right),u+\theta\phi,\xi\right]-f\left[x\left(u,\xi\right),u,\xi\right]}{\theta}$$

denotes the G-derivative of $f\left[x\left(u,t\right),u,t\right]$ relative to the direction ϕ. We assume that $f\left(\cdot,\cdot,\cdot\right)$ is continuously differentiable with respect to both x and u; hence, the variational chain rule gives

$$\delta x\left(u,\phi;t\right)=\int_{t_0}^{t}\left[\frac{\partial f\left[x\left(u,\xi\right),u,\xi\right]}{\partial x}\delta x\left(u,\phi;\xi\right)+\frac{\partial f\left[x\left(u,\xi\right),u,\xi\right]}{\partial u}\delta u\left(\phi\right)\right]d\xi \tag{4.21}$$

Observing that

$$\delta u\left(\phi\right)=\lim_{\theta\longrightarrow\mathbf{0}}\frac{\left(u+\theta\phi\right)-u}{\theta}=\phi$$

and employing the shorthand $y=\delta x\left(u,\phi;t\right)$, expression (4.21) becomes the integral equation

$$y=\int_{t_0}^{t_f}\left[\frac{\partial f}{\partial x}\cdot y+\frac{\partial f}{\partial u}\cdot\phi\right]dt \tag{4.22}$$

It is of course immediate from this integral equation that y obeys

$$\frac{dy}{dt}=\frac{\partial f}{\partial x}\cdot y+\frac{\partial f}{\partial u}\cdot\phi \tag{4.23}$$

$$y\left(t_0\right)=0 \tag{4.24}$$

which is recognized as an initial-value problem.

We now turn our attention to the G-derivative of J:

$$\delta J\left(u,\phi\right)=\lim_{\theta\longrightarrow\mathbf{0}_+}\frac{J\left(u+\theta\phi\right)-J\left(u\right)}{\theta} \tag{4.25}$$

To express this derivative we note that

$$\frac{J\left(u+\theta\phi\right)-J\left(u\right)}{\theta}=\int_{t_0}^{t_f}\frac{f_0\left(x\left(u+\theta\phi,t\right),u+\theta\phi,t\right)-f\left(x\left(u,t\right),u,t\right)}{\theta}dt$$

Using arguments completely analogous to those employed above in expressing y, we find that taking the limit of (4.25) as $\theta \longrightarrow 0_+$ leads to

$$\delta J\left(u,\phi\right)=\int_{t_0}^{t_f}\left[\frac{\partial f_0}{\partial x}\cdot y+\frac{\partial f_0}{\partial u}\cdot\phi\right]dt \tag{4.26}$$

where it is implicit that y depends on ϕ. We restate (4.26) by introducing *adjoint variables* λ defined by the final value problem

$$-\frac{d\lambda}{dt} = \left(\frac{\partial f}{\partial x}\right)^T \lambda + \left(\frac{\partial f_0}{\partial x}\right)^T \tag{4.27}$$

$$\lambda\left(t_f\right) = 0 \tag{4.28}$$

Note that there are n adjoint variables, one for each state equation. Furthermore, (4.27) is equivalent to

$$-\left(\frac{d\lambda}{dt}\right)^T - \lambda^T \frac{\partial f}{\partial x} = \frac{\partial f_0}{\partial x}$$

so that (4.26) becomes

$$\delta J\left(u, \phi\right) = \int_{t_0}^{t_f} \left[-\left(\frac{d\lambda}{dt}\right)^T y - \lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u}\phi\right] dt \tag{4.29}$$

Upon noting that

$$\left[\lambda^T y\right]_{t_0}^{t_f} = 0$$

since $\lambda\left(t_f\right) = 0$ and $y\left(t_0\right) = 0$, integration by parts combined with substitution based on (4.23) yields

$$\begin{aligned}\int_{t_0}^{t_f} -\left(\frac{d\lambda}{dt}\right)^T y dt &= \int_{t_0}^{t_f} \lambda^T \frac{dy}{dt} dt - \left[\lambda^T y\right]_{t_0}^{t_f} \\ &= \int_{t_0}^{t_f} \lambda^T \frac{dy}{dt} dt \\ &= \int_{t_0}^{t_f} \lambda^T \left[\frac{\partial f}{\partial x} \cdot y + \frac{\partial f}{\partial u} \cdot \phi\right] dt \end{aligned} \tag{4.30}$$

It follows that

$$\begin{aligned}\delta J\left(u, \phi\right) &= \int_{t_0}^{t_f} \left[-\left(\frac{d\lambda}{dt}\right)^T y - \lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u}\phi\right] dt \\ &= \int_{t_0}^{t_f} \left[\lambda^T \frac{\partial f}{\partial x} y + \lambda^T \frac{\partial f}{\partial u}\phi - \lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u}\phi\right] dt \\ &= \int_{t_0}^{t_f} \left[\lambda^T \frac{\partial f}{\partial u} + \frac{\partial f_0}{\partial u}\right] \phi dt \end{aligned} \tag{4.31}$$

when (4.30) is substituted into (4.29).

Expression (4.31) makes it clear that in the space $V = \left(L^2\left[t_0, t_f\right]\right)^n$ we have

$$\nabla_u J\left(u\right) = \lambda^T \frac{\partial f}{\partial u} + \frac{\partial f_0}{\partial u} \tag{4.32}$$

Furthermore, Theorem 4.13 tells us

$$\delta J\left(u^*, \phi\right) = 0 \quad \forall \phi \in \left(L^2\left[t_0, t_f\right]\right)^n$$

at an optimal control u^*. That is

$$\nabla_u J\left(u^*\right) = \lambda^T \frac{\partial f\left(x\left(u^*, t\right), u^*, t\right)}{\partial u} + \frac{\partial f_0\left(x\left(u^*, t\right), u^*, t\right)}{\partial u} = 0 \tag{4.33}$$

When the usual definition

$$H\left(x, u, \lambda, t\right) \equiv f_0\left(x, u, t\right) + \lambda f\left(x, u, t\right) \tag{4.34}$$

of the Hamiltonian is employed, (4.33) is readily seen to be the minimum principle; that is

$$u^* = \arg\min_u H\left(x, u, \lambda, t\right) \Longleftrightarrow \frac{\partial H}{\partial u} = 0 \quad \forall t \in \left[t_0, t_f\right]$$

since there are no control constraints. The remaining necessary conditions are the adjoint equation (4.27) and the transversality condition (4.28), plus of course the state initial-value problem, which is parametric in λ.

Thus, we have shown in this section that continuous-time optimal control necessary conditions, including the minimum principle, may be obtained directly from the theory of infinite-dimensional mathematical programming for the case of no terminal time constraints, no terminal costs, and no control constraints. In fact, it is possible to analyze the general continuous-time optimal control problem in a completely analogous fashion to obtain the same necessary conditions derived in Chapter 3 using the perspective of the classical calculus of variations (namely, the notion of a variation). This demonstration is quite important, for it establishes that there is indeed a single root problem (an infinite-dimensional mathematical program) that is the fundamental problem of dynamic optimization.

4.3.2 Variational Inequalities as Necessary Conditions

In this section, we want to consider a more general form of continuous-time optimal control and show its necessary conditions may be expressed as a variational inequality. To that end, we now employ the following form of the state operator:

$$x\left(u, t\right) = \arg\left\{\frac{dx}{dt} = f\left(x, u, t\right),\ x\left(t_0\right) = x_0,\ \Psi\left[x\left(t_f\right), t_f\right] = 0\right\} \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \tag{4.35}$$

arising in optimal control problem $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ with x_0, t_0, and t_f fixed; note that $x\left(t_f\right)$ obeys the terminal constraints intrinsic to (4.35). The relevant mappings are

$$
\begin{aligned}
f &: \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n \\
\Psi &: \Re^n \times \Re^1_+ \longrightarrow \Re^r \\
u &\in U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m
\end{aligned}
$$

where $\left(L^2\left[t_0, t_f\right]\right)^m$ is the m-fold product of the space of square-integrable functions $L^2\left[t_0, t_f\right]$, while $\left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ is the n-fold product of the Sobolev space $\mathcal{H}^1\left[t_0, t_f\right]$.

Additionally we invoke the following regularity condition for the two-point boundary-value problem (4.35):

Definition 4.37. *Regular state operator. We shall say the state operator $x(u,t)$ given by (4.35) is regular if it exists for all admissible u and is unique, strongly continuous and G-differentiable with respect to u.*

The notation $x(u,t)$ is of course familiar from the discussion of Section 4.3.1; it denotes an operator which determines the state vector for each control vector. In order to use the state operator notation $x(u,t)$, we will invoke regularity in the sense of Definition 4.37 to ensure that the two-point boundary-value problem of (4.35) is well posed. In particular, the two-point boundary-value problem of (4.35) is assumed to have a unique solution for all admissible controls. In the event a terminal state $x\left(t_f\right)$ fulfilling the terminal constraints $\Psi\left[x\left(t_f\right), t_f\right] = 0$ cannot be reached from the intial state x_0 for controls $u \in U^\dagger \subset U$, then the control constraints must be restated as

$$
u \in U \backslash U^\dagger \subseteq \left(L^2\left[t_0, t_f\right]\right)^m \tag{4.36}
$$

In the material that follows, we assume any such restatement to assure reachability has already occurred and regularity in the sense of Definition 4.37 holds.

Keeping in mind that t_0, t_f, and x_0 are fixed, the optimal control problem we consider is

$$
\min J(u) = K\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} f_0(x,u,t)dt \tag{4.37}
$$

subject to

$$
\frac{dx}{dt} = f(x,u,t) \tag{4.38}
$$

$$
x(t_0) = x_0 \in \Re^n \tag{4.39}
$$

$$
u \in U \tag{4.40}
$$

$$
\Psi\left[x\left(t_f\right), t_f\right] = 0 \tag{4.41}
$$

where

$$K : \Re^n \times \Re^1_+ \longrightarrow \Re^1$$
$$f_0 : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow L^2\left[t_0, t_f\right]$$

We wish to formally state and prove the following result:

Theorem 4.16. *Necessary conditions for optimal control. Consider the optimal control problem OCP*$(f_0, f, K, \Psi, U, x_0, t_0, t_f)$ *defined by (4.37), (4.38), (4.39), (4.40), and (4.41) with* t_0, $x(t_0)$, *and* t_f *fixed. Suppose the following regularity conditions are satisfied*

R1. $u \in U \subseteq \left(L^2\left[t_0, \tau\right]\right)^m$
R2. *the operator* $x(u,t) : \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ *is regular in the sense of Definition 4.37;*
R3. $K : \Re^n \times \Re^1_+ \longrightarrow \Re^1$ *is continuously differentiable with respect to* x *and* t*;*
R4. $\Psi : \Re^n \times \Re^1_+ \longrightarrow \Re^r$ *is continuously differentiable with respect to* x *and* t*;*
R5. $f_0 : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow L^2\left[t_0, t_f\right]$ *is continuously differentiable with respect to* x *and* u*;*
R6. $f : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n$ *is continuously differentiable with respect to* x *and* u*;*
R7. $x_0 \in \Re^n$, $t_0 \in \Re^1_+$, *and* $t_f \in \Re^1_{++}$ *are known and fixed;*
R8. U *is convex; and*
R9. $\upsilon \in \Re^r$.

Then any solution $u^* \in U$ *obeys the following necessary conditions:*

N1. *the variational inequality:*

$$\sum_{i=1}^{m} \frac{\partial H\left(x^*, u^*, \lambda^*, t\right)}{\partial u_i}\left(u_i - u_i^*\right) \geq 0 \quad \forall u \in U \tag{4.42}$$

where

$$H(x, u, \lambda, t) = f_0(x, u, t) + \lambda^T f(x, u, t)$$

N2. *the state initial-value problem:*

$$\frac{dx^*}{dt} = f\left(x^*, u^*, t\right)$$
$$x^*(t_0) = x_0;$$

N3. *the adjoint dynamics:*

$$(-1)\frac{d\lambda^*}{dt} = \nabla_x H^*; \quad \textit{and}$$

N4. *the transversality conditions:*

$$\lambda\left(t_f\right) = \frac{\partial \Phi\left[x^*\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)}$$
$$\Phi\left[x^*\left(t_f\right), t_f\right] = K\left[x^*\left(t_f\right), t_f\right] + \upsilon^T \Psi\left[x^*\left(t_f\right), t_f\right]$$

Proof. Note that

$$x(u,t) = x(t_0) + \int_{t_0}^{t} f\left[x(u,\xi), u, \xi\right] d\xi$$

It is immediate that

$$x(u+\theta\rho,t) = x(t_0) + \int_{t_0}^{t} f\left[x(u+\theta\rho), u+\theta\rho, \xi\right] d\xi$$

Consequently,

$$\delta x(u,\rho;t) = \int_{t_0}^{t} \left\{ \frac{\partial f\left[x(u,\xi), u, \xi\right]}{\partial x} \delta x(u,\rho;t) + \frac{\partial f\left[x(u,\xi), u, \xi\right]}{\partial u} \delta u(\rho) \right\} d\xi$$

where of course the G-derivative of u obeys

$$\delta u(\rho) = \lim_{\theta \longrightarrow 0} \frac{(u+\theta\rho) - u}{\theta} = \rho$$

Employing the shorthand $y = \delta x(u,\rho;t)$, we have the integral equation

$$y = \int_{t_0}^{t} \left[\frac{\partial f}{\partial x} y + \frac{\partial f}{\partial u} \rho \right] dt \tag{4.43}$$

It is of course immediate from this integral equation that y obeys

$$\frac{dy}{dt} = \frac{\partial f}{\partial x} y + \frac{\partial f}{\partial u} \rho; \quad y(t_0) = 0 \tag{4.44}$$

which is recognized as an initial value problem, verifying that the G-derivative of x is well defined. The G-derivative of J obeys

$$\begin{aligned} \delta J(u,\rho) &= \left[\frac{\partial \Phi\left[x(t), t\right]}{\partial x} \delta x(u,\rho;t) \right]_{t=t_f} \\ &\quad + \int_{t_0}^{t_f} \left[\frac{\partial f_0}{\partial x} \delta x(u,\rho;t) + \frac{\partial f_0}{\partial u} \delta u(\rho) \right] \\ &= \frac{\partial \Phi\left[x\left(t_f\right), t_f\right]}{\partial x} y\left(t_f\right) + \int_{t_0}^{t_f} \left[\frac{\partial f_0}{\partial x} y + \frac{\partial f_0}{\partial u} \rho \right] dt \end{aligned}$$

We introduce *adjoint variables* λ defined by the following final value problem

$$-\frac{d\lambda}{dt} = \left(\frac{\partial f}{\partial x}\right)^T \lambda + \left(\frac{\partial f_0}{\partial x}\right)^T; \ \lambda\left(t_f\right) = \frac{\partial \Phi\left[x\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)} \tag{4.45}$$

so that

$$\delta J\left(u, \rho\right) = \int_{t_0}^{t_f} \left[-\left(\frac{d\lambda}{dt}\right)^T y - \lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u}\rho\right] dt \tag{4.46}$$

Note that

$$\left[\lambda^T y\right]_{t_0}^{t_f} = \left[\lambda\left(t_f\right)\right]^T y\left(t_f\right) - \left[\lambda\left(t_0\right)\right]^T y\left(t_0\right) = \frac{\partial \Psi\left[x\left(t_f\right), t_f\right]}{\partial x} y\left(t_f\right)$$

due to (4.45) and the fact that $y\left(t_0\right) = 0$, so an integration by parts yields

$$\begin{aligned}
\int_{t_0}^{t_f} -\left(\frac{d\lambda}{dt}\right)^T y dt &= \int_{t_0}^{t_f} \lambda^T \frac{dy}{dt} dt - \left[\lambda^T y\right]_{t_0}^{t_f} \\
&= \int_{t_0}^{t_f} \lambda^T \frac{dy}{dt} dt - \frac{\partial \Phi\left[x\left(t_f\right), t_f\right]}{\partial x} y\left(t_f\right) \\
&= \int_{t_0}^{t_f} \lambda^T \left[\frac{\partial f}{\partial x} \cdot y + \frac{\partial f}{\partial u} \cdot \rho\right] dt - \frac{\partial \Phi\left[x\left(t_f\right), t_f\right]}{\partial x} y\left(t_f\right)
\end{aligned} \tag{4.47}$$

It follows that

$$\begin{aligned}
\delta J\left(u, \rho\right) &= \frac{\partial \Phi\left[x\left(t_f\right), t_f\right]}{\partial x} y\left(t_f\right) + \int_{t_0}^{t_f} \left\{\lambda^T \left[\frac{\partial f}{\partial x} \cdot y + \frac{\partial f}{\partial u} \cdot \rho\right]\right. \\
&\quad \left. -\lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u}\rho\right\} dt - \frac{\partial \Phi\left[x\left(t_f\right), t_f\right]}{\partial x} y\left(t_f\right) \\
&= \int_{t_0}^{t_f} \left[\lambda^T \frac{\partial f}{\partial u} + \frac{\partial f_0}{\partial u}\right] \rho dt
\end{aligned}$$

By virtue of the definition of the Hamiltonian, we have

$$\delta J\left(u, \rho\right) = \int_{t_0}^{t_f} \left[\frac{\partial H}{\partial u}\rho\right] dt \tag{4.48}$$

as an expression for the G-derivative of the criterion with respect to u. By Theorem 4.15, optimality requires $u^* \in U$ to obey

$$\delta J(u^*, \rho) \geq 0 \quad \forall \rho = u - u^* \tag{4.49}$$

Thus, by statement (4.16), we know

$$\delta J(u^*, u - u^*) = \langle \nabla J(u^*), u - u^* \rangle \geq 0 \quad \forall u \in U \tag{4.50}$$

where

$$[\nabla J\ (u)]_i = \frac{\partial H}{\partial u_i} \tag{4.51}$$

for $t \in \left[t_0, t_f\right]$. The desired necessary conditions (4.42) are immediate from (4.50) and (4.51). ■

It is tempting to say that, if one additionally requires the Hamiltonian to be convex in its controls u, then the necessary conditions of Theorem 4.16 above will become sufficient. This is, in general, not true. Rather, merely imposing the requirement for the Hamiltonian to be convex in u will only make the variational inequality (4.42) equivalent to the minmum principle. To assure sufficiency of the necessary conditions of Theorem 4.17, one must require the Hamiltonian to be convex in (x, u), thereby allowing application of an extended version of the Mangasarian sufficiency theorem presented in Chapter 3.

It should also be clear to the reader that the variational inequality (4.42), which holds for each instant of time, leads directly to the following infinite-dimensional variational inequality:

$$\int_{t_0}^{t_f} \sum_{i=1}^{m} \frac{\partial H\left[x\left(u^*, t\right), u^*, \lambda\left(u^*, t\right), t\right]}{\partial u_i} \left(u_i - u_i^*\right) dt \geq 0 \quad \forall u \in U \tag{4.52}$$

which is also a necessary condition. In (4.52), we have taken some notational liberties in introducing the adjoint operator $\lambda\ (u, t)$ without having formally defined it; hopefully this informality does not hamper the reader's understanding. We will have more to say about infinite-dimensional variational inequalies in Chapter 6.

4.4 Optimal Control with Time Shifts

It is possible to extend the results obtained previously for deterministic optimal control to consider the possibility that control variables with state-dependent time shifts may appear in the formulation of some optimal control problem of interest. In particular, we will allow both the integrand of the criterion functional and the dynamics themselves to involve time shifts that are state-dependent. Such time-shifted problems arise in dynamic traffic assignment, logistics, and supply chain modeling, as well as in other applications.

4.4.1 Some Preliminaries

We now consider a somewhat different state operator

$$x(u, u_\tau, t) = \arg\left\{ \frac{dx}{dt} = f(x, u, u_\tau, t),\ x(t_0) = x_0, \Psi\left[x\left(t_f\right), t_f\right] = 0 \right\} \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \tag{4.53}$$

where t_0, t_f, and x_0 are fixed and

$$\left[t_0, t_f\right] \subset \Re^1_+$$

Furthermore $u_\tau(t)$ is a shorthand for the shifted control vector

$$u_\tau(t) = \begin{pmatrix} u_1(t - \tau_1(x)) \\ \vdots \\ u_m(t - \tau_m(x)) \end{pmatrix}$$

where

$$\tau_i : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \longrightarrow \Re^1_+$$

for each $i \in [1, m]$. Other relevant mappings needed in this discussion are

$$\begin{aligned} f &: \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n \\ \Psi &: \Re^n \times \Re^1_+ \longrightarrow \Re^r \\ K &: \Re^n \times \Re^1_+ \longrightarrow \Re^1 \\ u &\in U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m \\ u_\tau &\in \left(L^2\left[t_0, t_f\right]\right)^m \end{aligned}$$

and

$$\Phi\left[x\left(t_f\right), t_f, v\right] = K\left[x\left(t_f\right), t_f\right] + v^T \Psi\left[x\left(t_f\right), t_f\right] \tag{4.54}$$

As is by now familiar, $\left(L^2\left[t_0, t_f\right]\right)^m$ is the m-fold product of the space of square-integrable functions $L^2\left[t_0, t_f\right]$, while $\left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ is the n-fold product of the Sobolev space $\mathcal{H}^1\left[t_0, t_f\right]$. We invoke the following regularity condition for the time-shifted two-point boundary-value problem intrinsic to (4.53):

Definition 4.38. *Regular dynamics. We shall say the state operator $x(u, u_\tau, t)$ given by (4.53) is regular if it exists for all admissible u and is unique, strongly continuous, and G-differentiable with respect to u and u_τ.*

Remarks about reachability analogous to those made following Definition 4.37 apply to Definition 4.38 and the interpretation of $x(u, u_\tau, t)$.

4.4.2 The Optimal Control Problem of Interest

Keeping in mind t_0, t_f, and x_0 are fixed, we consider the following optimal control problem:

$$\min K\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} f_0(x, u, u_\tau, t)dt \tag{4.55}$$

subject to

$$\frac{dx}{dt} = f\left(x, u, u_\tau, t\right) \tag{4.56}$$

$$x\left(t_0\right) = x_0 \tag{4.57}$$

$$u \in U \tag{4.58}$$

$$\Psi\left[x\left(t_f\right), t_f\right] = 0 \tag{4.59}$$

This is a nonstandard optimal control problem due to the presence of state-dependent time shifts, and we will need to derive its necessary conditions. We refer to this optimal control problem as $OCP(f_0, f, K, \Psi, U, x_0, t_0, t_f, \tau)$.

4.4.3 Change of Variable

Let us briefly consider the following abstract integral:

$$I = \int_{t_0}^{t_f} F(t)dt$$

For this integral we wish to make a change of variables, namely

$$s = t - w(t)$$

Thus, we have

$$ds = \left(1 - \frac{dw}{dt}\right)dt$$

or equivalently

$$dt = \frac{1}{1 - \dot{w}}ds$$

Therefore, we have

$$I = \int_{t_0 - w(t_0)}^{t_f - w(t_f)} \left[\frac{F(t)}{1 - \dot{w}\left(t\right)}\right]_{t = s + w(t)} ds \tag{4.60}$$

The meaning of

$$\left[\frac{F(t)}{1 - \dot{w}\left(t\right)}\right]_{t = s + w(t)}$$

is that the integrand of (4.60) is to be evaluated at the instant of time t that solves the fixed-point problem

$$t = s + w(t)$$

for each $s \in \left[t_0 - w(t_0), t_f - w(t_f)\right]$ that is visited during the integration process. Note that the roles of s and t may be exchanged, so that (4.60) becomes the equivalent expression

$$I = \int_{t_0 - w(t_0)}^{t_f - w(t_f)} \left[\frac{F(s)}{1 - \dot{w}(s)}\right]_{s=t+w(s)} dt$$

In developing necessary conditions for optimal control with time shifts, we will employ a change of variables and related notation similar to that introduced above.

4.4.4 Necessary Conditions for Time-Shifted Problems

We will state and prove the following result that extends the analysis by Budelis and Bryson (1970) of problems involving time shifts that are fixed:

Theorem 4.17. *Necessary conditions for optimal control with time shifts. Consider the optimal control problem OCP*$(f_0, f, K, \Psi, U, x_0, t_0, t_f, \tau)$ *defined by (4.55), (4.56), (4.57), (4.58), and (4.59) with* t_0, $x(t_0)$, *and* t_f *fixed. Suppose the following regularity conditions are satisfied:*

R1. $u \in U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m$;
R2. $u_\tau \in U \in \left(L^2\left[t_0, t_f\right]\right)^m$;
R3. $x(u, u_\tau, t) : \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ *is regular in the sense of Definition 4.38*;
R4. τ *is continuously differentiable with respect to* x;
R5. $K : \Re^n \times \Re^1_+ \longrightarrow \Re^1$ is continuously differentiable with respect to x and t;
R6. $\Psi : \Re^n \times \Re^1_+ \longrightarrow \Re^r$ is continuously differentiable with respect to x and t;
R7. $f_0 : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow L^2\left[t_0, t_f\right]$ *is continuously differentiable with respect to* x, u, *and* u_τ;
R8. $f : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n$ *is continuously differentiable with respect to* x, u, *and* u_τ;
R9. $x_0 \in \Re^n$, $t_0 \in \Re^1_+$, *and* $t_f \in \Re^1_{++}$ *are known and fixed;*
R10. U is convex; and
R11. $v \in \Re^r$.

Then any solution $u^* \in U$ *obeys the following necessary conditions:*

N1. *the variational inequality:*

$$\sum_{i=1}^{m} \frac{\partial H_1\left(x^*, u^*, u^*_\tau, \lambda^*, t\right)}{\partial u_i}\left(u_i - u^*_i\right) \geq 0 \quad \forall u \in U$$

where

$$H_1(x,u,u_\tau,\lambda,t) = f_0(x,u,u_\tau,t) + \lambda^T f(x,u,u_\tau,t);$$

N2. *the state dynamics:*

$$\frac{dx^*}{dt} = f\left(x^*,u^*,u_\tau^*,t\right)$$
$$x^*(t_0) = x_0;$$

N3. *the adjoint dynamics:*

$$(-1)\frac{d\lambda^*}{dt} = \nabla_x\left(\lambda^*\right)^T f\left(x^*,u^*,u_\tau^*,t\right); \text{ and}$$

N4. *the transversality condition:*

$$\lambda^*\left(t_f\right) = \frac{\partial\Phi\left[x^*\left(t_f\right),t_f,v\right]}{\partial x}$$

Proof. A similar result was proven by Budelis and Bryson (1970). Our proof differs from theirs by making the time-shifts state-dependent, relying on the notion of a G-derivative, and emphasizing the variational inequality version of the necessary conditions for optimal control that are the subject of Theorem 4.15. We begin by noting that

$$x(u,u_\tau,t) = x(t_0) + \int_{t_0}^{t} f\left[x(u,u_\tau,\xi),u,u_\tau,\xi\right]d\xi$$

It is immediate that

$$x(u+\theta\rho,u_\tau+\theta\rho_\tau,t) =$$
$$x(t_0) + \int_{t_0}^{t} f\left[x(u+\theta\rho,u_\tau+\theta\rho_\tau,\xi),u+\theta\rho,u_\tau+\theta\rho_\tau,\xi\right]d\xi$$

Consequently,

$$\delta x(u,\rho;u_\tau,\rho_\tau;t) = \int_{t_0}^{t} \frac{\partial f\left[x(u,u_\tau,\xi),u,u_\tau,\xi\right]}{\partial x}\delta x(u,\rho;u_\tau,\rho_\tau;t)$$
$$+\frac{\partial f\left[x(u),u,u_\tau,\xi\right]}{\partial u}\delta u(\rho)$$
$$+\frac{\partial f\left[x(u),u,u_\tau,\xi\right]}{\partial u_\tau}\delta u_\tau(\rho_\tau)\,d\xi$$

where the G-derivatives of u and u_τ obey

$$\delta u(\rho) = \lim_{\theta \longrightarrow 0} \frac{(u + \theta\rho) - u}{\theta} = \rho \quad \delta u_\tau(\rho_\tau) = \lim_{\theta \longrightarrow 0} \frac{(u_\tau + \theta\rho_\tau) - u_\tau}{\theta} = \rho_\tau \tag{4.61}$$

Employing the shorthand $y = \delta x(u, \rho; u_\tau, \rho_\tau; t)$, we have the integral equation

$$y = \int_{t_0}^{t} \left[\frac{\partial f}{\partial x} y + \frac{\partial f}{\partial u} \rho + \frac{\partial f}{\partial u_\tau} \rho_\tau \right] dt \tag{4.62}$$

It is of course immediate from this integral equation that y obeys

$$\frac{dy}{dt} = \frac{\partial f}{\partial x} y + \frac{\partial f}{\partial u} \rho + \frac{\partial f}{\partial u_\tau} \rho_\tau; \quad y(t_0) = 0 \tag{4.63}$$

which is recognized as an initial value problem, verifying that the G-derivative of x is well defined. The G-derivative of J obeys

$$\begin{aligned} \delta J(u, \rho; u_\tau, \rho_\tau) = & \left[\frac{\partial \Phi[x(t), t, v]}{\partial x} \delta x(u, \rho; u_\tau, \rho_\tau) \right]_{t_0}^{t_f} \\ & + \int_{t_0}^{t_f} \left[\frac{\partial f_0}{\partial x} \delta x(u, \rho; u_\tau, \rho_\tau) + \frac{\partial f_0}{\partial u} \delta u(\rho) + \frac{\partial f_0}{\partial u_\tau} \delta u(\rho_\tau) \right] \\ = & \frac{\partial \Phi[x(t_f), t_f, v]}{\partial x} y(t_f) + \int_{t_0}^{t_f} \left[\frac{\partial f_0}{\partial x} y + \frac{\partial f_0}{\partial u} \rho + \frac{\partial f_0}{\partial u_\tau} \rho_\tau \right] dt \end{aligned}$$

We introduce *adjoint variables* λ defined by the final value problem

$$-\frac{d\lambda}{dt} = \left(\frac{\partial f}{\partial x} \right)^T \lambda + \left(\frac{\partial f_0}{\partial x} \right)^T; \quad \lambda(t_f) = \frac{\partial \Phi[x(t_f), t_f, v]}{\partial x} \tag{4.64}$$

so that

$$\begin{aligned} \delta J(u, \rho; u_\tau, \rho_\tau) = & \frac{\partial \Phi[x(t_f), t_f, v]}{\partial x} y(t_f) \\ & + \int_{t_0}^{t_f} \left[-\left(\frac{d\lambda}{dt} \right)^T y - \lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u} \rho + \frac{\partial f_0}{\partial u_\tau} \rho_\tau \right] dt \end{aligned}$$

Note that

$$\left[\lambda^T y \right]_{t_0}^{t_f} = \left[\lambda(t_f) \right]^T y(t_f) - \left[\lambda(t_0) \right]^T y(t_0) = \frac{\partial \Phi[x(t_f), t_f, v]}{\partial x} y(t_f)$$

due to (4.64) and the fact that $y(t_0) = 0$, so an integration by parts yields

$$\begin{aligned}\int_{t_0}^{t_f} -\left(\frac{d\lambda}{dt}\right)^T y dt &= \int_{t_0}^{t_f} \lambda^T \frac{dy}{dt} dt - \left[\lambda^T y\right]_{t_0}^{t_f} \\ &= \int_{t_0}^{t_f} \lambda^T \frac{dy}{dt} dt - \frac{\partial \Phi\left[x\left(t_f\right), t_f, v\right]}{\partial x} y\left(t_f\right) \\ &= \int_{t_0}^{t_f} \lambda^T \left[\frac{\partial f}{\partial x} \cdot y + \frac{\partial f}{\partial u} \cdot \rho + \frac{\partial f}{\partial u_\tau} \rho_\tau\right] dt \\ &\quad - \frac{\partial \Phi\left[x\left(t_f\right), t_f, v\right]}{\partial x} y\left(t_f\right)\end{aligned}$$

It follows that

$$\begin{aligned}\delta J(u, \rho; u_\tau, \rho_\tau) &= \frac{\partial \Phi\left[x\left(t_f\right), t_f, v\right]}{\partial x} y\left(t_f\right) \\ &\quad + \int_{t_0}^{t_f} \left\{\lambda^T \left[\frac{\partial f}{\partial x} \cdot y + \frac{\partial f}{\partial u} \cdot \rho + \frac{\partial f}{\partial u_\tau} \rho_\tau\right]\right. \\ &\quad \left. -\lambda^T \frac{\partial f}{\partial x} y + \frac{\partial f_0}{\partial u} \rho + \frac{\partial f_0}{\partial u_\tau} \rho_\tau\right\} dt - \frac{\partial \Phi\left[x\left(t_f\right), t_f, v\right]}{\partial x} y\left(t_f\right) \\ &= \int_{t_0}^{t_f} \left[\lambda^T \frac{\partial f}{\partial u} + \frac{\partial f_0}{\partial u}\right] \rho dt + \int_{t_0}^{t_f} \left[\lambda^T \frac{\partial f}{\partial u_\tau} + \frac{\partial f_0}{\partial u_\tau}\right] \rho_\tau dt\end{aligned}$$

Since $H_1(x, u, u_\tau, \lambda, t) = f_0(x, u, u_\tau, t) + \lambda^T f(x, u, u_\tau, t)$, we have

$$\delta J(u, \rho; u_\tau, \rho_\tau) = \int_{t_0}^{t_f} \left[\frac{\partial H_1}{\partial u} \rho + \frac{\partial H_1}{\partial u_\tau} \rho_\tau\right] dt \tag{4.65}$$

as an expression for the G-derivative of the criterion with respect to both u and u_τ. Moreover, terms of the form

$$\int_{t_0}^{t_f} \frac{\partial H_1}{\partial (u_\tau)_i} (\delta u_\tau)_i \, dt = \int_{t_0}^{t_f} \frac{\partial H_1}{\partial (u_\tau)_i} \delta u_i \left(t + \tau_i(x_i)\right) dt$$

may be re-expressed by making the change of variables

$$t = s_i + \tau_i[x(t)]$$

Because the $\tau_i\,(x)$ are differentiable with respect to x_i, the implicit function theorem gives

$$\frac{dt}{ds_i} = -\frac{\partial\left[t - s_i - \tau_i\,(x)\right]/\partial s_i}{\partial\left[t - s_i - \tau_i\,(x)\right]/\partial t} = \frac{1}{1 - \sum_{j=1}^{n} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j}$$

or,

$$dt = \frac{1}{1 - \sum_{j=1}^{n} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j} ds_i \tag{4.66}$$

Note that

$$t = t_0 \Longrightarrow s_i = t_0 - \tau_i\left[x\,(t_0)\right];\; t = t_f \Longrightarrow s_i = t_f - \tau_i\left[x\left(t_f\right)\right]$$

A change of variables based on (4.66) leads to

$$\int_{t_0}^{t_f} \frac{\partial H_1}{\partial\,(u_\tau)_i}\delta\,(u_\tau)_i\,dt = \int_{t_0-\tau_i(x(t_0))}^{t_f-\tau_i\left(x_i\left(t_f\right)\right)} \left[\frac{\partial H_1}{\partial\,(u_\tau)_i}\frac{1}{1 - \sum_{j=1}^{n} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j}\right]_{s_i} \delta\,(u)_i\,dt \tag{4.67}$$

where $s_i\,(t)$ obeys $s_i\,(t) = \arg\left[s = t + \tau_i\left[x\,(s)\right]\right]$ for any time t visited during the integration process; that is, the expression

$$\frac{\partial H_1}{\partial\,(u_\tau)_i}\frac{1}{1 - \sum_{j=1}^{n} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j}$$

taken from (4.67) is meant to be evaluated at s_i. Furthermore, without loss of generality, we may consider $\delta u_i = 0$ for any time $t < t_0$. Therefore, upon recalling (4.61), we have from (4.67) the following:

$$\int_{t_0}^{t_f} \frac{\partial H_1}{\partial\,(u_\tau)_i}\delta\,(u_\tau)_i\,dt = \int_{t_0}^{t_f-\tau_i\left(x_i\left(t_f\right)\right)} \left[\frac{\partial H_1}{\partial\,(u_\tau)_i}\frac{1}{1 - \sum_{j=1}^{n} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j}\right]_{s_i(t)} \rho_i\,dt \tag{4.68}$$

since $\delta\,(u) = \rho$. We next note that

$$\int_{t_0}^{t_f} \frac{\partial H_1}{\partial u_i}\rho_i\,dt = \int_{t_0}^{t_f-\tau_i\left(x_i\left(t_f\right)\right)} \frac{\partial H_1}{\partial u_i}\rho_i\,dt + \int_{t_f-\tau_i\left(x_i\left(t_f\right)\right)}^{t_f} \frac{\partial H_1}{\partial u_i}\rho_i\,dt \tag{4.69}$$

From (4.65), (4.68), and (4.69) we see that

$$
\begin{aligned}
[\delta J\,(u,\rho;u_\tau,\rho_\tau)]_i &= [\delta J\,(u,\rho)]_i \\
&\equiv \int_{t_f-\tau_i(x_i(t_f))}^{t_f} \frac{\partial H_1}{\partial u_i}\rho_i\,dt \\
&\quad + \int_{t_0}^{t_f-\tau_i(x(t_f))} \left\{ \frac{\partial H_1}{\partial u_i} + \left[\frac{\partial H_1}{\partial\,(u_\tau)_i} \frac{1}{1-\sum_{j=1}^{n} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j} \right]_{s_i} \right\} \rho_i\,dt
\end{aligned}
$$

which establshes that

$$
[\nabla J\,(u)]_i = \frac{\partial H_1}{\partial u_i} \quad \text{if } t \in \left[t_f - \tau_i\left(x\left(t_f\right)\right), t_f\right]
$$

$$
[\nabla J\,(u)]_i = \frac{\partial H_1}{\partial u_i} + \left[\frac{\partial H_1}{\partial\,(u_\tau)_i} \frac{1}{1-\sum_{j=1}^{m} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j} \right]_{s_i} \quad \text{if } t \in \left[t_0, t_f-\tau_i\left[x\left(t_f\right)\right]\right]
$$

The variational inequality optimality condition

$$
\delta J\left(u^*,\rho\right) = \left\langle \nabla J\left(u^*\right),\rho\right\rangle = \sum_{i=1}^{m} \int_{t_0}^{t_f} \left[\nabla J\left(u^*\right)\right]_i \left(u_i - u_i^*\right) dt \geq 0 \quad \forall u \in U
$$

directly yields the desired variational inequality necessary condition when it is observed that each direction may be stated as $\rho = (u - u^*)$ for $u \in U$. ■

The following result, stemming directly from the final remarks of the proof imediately above, is important:

Corollary 4.1. *Gradient of the criterion in the presence of time shifts. Under the given of Theorem 4.17, the gradient of the criterion (4.55) is defined by*

$$
[\nabla J\,(u)]_i = \frac{\partial H_1}{\partial u_i} \quad \textit{if } t \in \left[t_f - \tau_i\left(x\left(t_f\right)\right), t_f\right]
$$

$$
[\nabla J\,(u)]_i = \frac{\partial H_1}{\partial u_i} + \left[\frac{\partial H_1}{\partial\,(u_\tau)_i} \frac{1}{1-\sum_{j=1}^{m} \dfrac{\partial \tau_i\,(x)}{\partial x_j}\dot{x}_j} \right]_{s_i(t)} \quad \textit{if } t \in \left[t_0, t_f-\tau_i\left(x\left(t_f\right)\right)\right]
$$

for $i = [1,m]$.

4.4.5 A Simple Abstract Example

Consider the optimal control problem

$$\min J(u) = \int_0^1 \frac{1}{2}u^2 dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= x + u(t-\psi) \\ 0 &\le u \le 1 \\ x(0) &= 1 \end{aligned}$$

where ψ is a constant scalar time lag obeying

$$0 < \psi < 1$$

and we use the abbreviation

$$u_\psi \equiv u(t-\psi)$$

to denote the time shifted control. By inspection, we see that the null control ($u^* = 0$) will be optimal for this problem.

We know the relevant Hamiltonian is

$$H = \frac{1}{2}u^2 + \lambda [x + u(t-\psi)]$$

We will need the partial derivatives

$$\begin{aligned} \frac{\partial H}{\partial u} &= u \\ \frac{\partial H}{\partial u_\psi} &= \lambda \\ \frac{\partial H}{\partial x} &= \lambda \end{aligned}$$

Let us first consider $t \in [1-\psi, 1]$. The optimal control problem is one for which the variational inequality principle is

$$\frac{\partial H^*}{\partial u}\left(u - u^*\right) = u^*\left(u - u^*\right) \ge 0 \qquad 0 \le u \le 1 \text{ and } 0 \le u^* \le 1 \quad t \in [1-\psi, 1] \tag{4.70}$$

Any solution of (4.70) must solve the following mathematical program

$$\min u^* u \quad \text{s.t.} \quad 1 - u \le 0 \text{ and } -u \le 0$$

which has the solution

$$u^* = 0 \quad \text{for } t \in [1-\psi, 1]$$

Let us now consider $t \in [0, 1-\psi]$. An optimal solution u^* must obey the variational inequality principle

$$\left\{ \frac{\partial H^*}{\partial u} + \left[\frac{\partial H^*}{\partial \left(u_\psi\right)_i} \frac{1}{1 - \dfrac{d\psi}{dt}} \right]_{s(t)} \right\} \left(u - u^*\right) \geq 0$$

Therefore, u^* solves

$$\left\{u^* + [\lambda]_{s(t)}\right\} \left(u - u^*\right) \geq 0 \quad 0 \leq u \leq 1 \text{ and } 0 \leq u^* \leq 1 \quad t \in [0, 1-\psi]$$

where

$$s(t) = t + \psi$$

Furthermore

$$(-1)\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = \lambda$$
$$\lambda(1) = 0$$

Consequently

$$\lambda = K \exp(-t)$$

and

$$\lambda(1) = \frac{K}{e} = 0 \Longrightarrow K = 0 \Longrightarrow \lambda = 0 \text{ for } t \in [0, 1-\psi]$$

Furthermore, for this unusually easy example problem, we have

$$[\lambda]_{s(t)} = \lambda(s) = \lambda(t + \psi) = 0$$

Thus, the relevant variational inequality is

$$\left(u^* + 0\right)\left(u - u^*\right) \geq 0 \qquad 0 \leq u \leq 1 \text{ and } 0 \leq u^* \leq 1 \quad t \in [0, 1-\psi] \quad (4.71)$$

It is trivial to show via the Kuhn-Tucker conditions for (4.71) that again

$$u^* = 0 \quad \text{for } t \in [0, 1-\psi]$$

Consequently

$$u^* = 0 \quad \text{for } t \in [0, 1]$$

as expected.

4.5 Derivation of the Euler-Lagrange Equation

In this section we are concerned with the following problem:

$$\min J(x) = \int_{t_0}^{t_f} f_0[x(t), \dot{x}(t), t]\, dt \tag{4.72}$$

$$x(t_0) = x_0 \tag{4.73}$$

$$x(t_f) = x_f \tag{4.74}$$

where x_0 and x_f are known (fixed) vectors; t_0 and t_f are also fixed. The functional $J(x)$ is to be minimized on $\left(C^1[t_0, t_f]\right)^n$, the vector space of continuous functions with continuous first derivatives relative to the segment $[t_0, t_f]$ of the nonnegative real line $\Re_+^1$. This is the foundation problem of the classical calculus of variations, that we studied in Chapter 3. It can be shown that $C^1[t_0, t_f]$ is not complete and, hence, not a Banach space nor a Hilbert space. The implication of $C^1[t_0, t_f]$ not being a Hilbert space is that the Riesz representation theorem does not hold and the G-derivative $\delta J(x, \phi)$ is not a continuous linear form from which we can identify the gradient of the functional $J(x)$.

The relevant G-derivative for directions $\phi \in \left(C^1[t_0, t_f]\right)^n$ is

$$\begin{aligned}\delta J(x, \phi) &= \lim_{\theta \longrightarrow 0} \frac{J(x + \theta\phi) - J(x)}{\theta} \\ &= \left[\frac{d}{d\theta} \int_{t_0}^{t_f} f_0\left(x + \theta\phi, \dot{x} + \theta\dot{\phi}, t\right) dt\right]_{\theta=0}\end{aligned} \tag{4.75}$$

If we make the definition

$$I \equiv \int_{t_0}^{t_f} f_0\left(x + \theta\phi, \dot{x} + \theta\dot{\phi}, t\right) dt$$

then the chain rule tells us that for all $\phi \in C^1[t_0, t_f]$

$$\begin{aligned}\delta J(x, \phi) &= \left[\frac{\partial I}{\partial(x + \theta\phi)} \frac{\partial(x + \theta\phi)}{\partial\theta} + \frac{\partial I}{\partial\left(\dot{x} + \theta\dot{\phi}\right)} \frac{\partial\left(\dot{x} + \theta\dot{\phi}\right)}{\partial\theta}\right]_{\theta=0} \\ &= \int_{t_0}^{t_f} \left[\frac{\partial f_0(x, \dot{x}, t)}{\partial x}\phi + \frac{\partial f_0(x, \dot{x}, t)}{\partial \dot{x}}\dot{\phi}\right] dt\end{aligned} \tag{4.76}$$

where the partial derivatives are to be interpreted as (transposed) gradients according to the usual scheme. From the preceding discussion, it is clear that the gradient of

J cannot be found in the usual way. Furthermore, because the endpoints x_0 and x_f are fixed, we expect that

$$\delta J\left[x\left(t_0\right), \phi\left(t_0\right)\right] = \lim_{\theta \longrightarrow 0} \frac{J\left[x\left(t_0\right) + \theta\phi\left(t_0\right)\right] - J\left[x\left(t_0\right)\right]}{\theta} = 0 \tag{4.77}$$

a fact that is ensured by taking $\phi\left(t_0\right) = 0$; similar reasoning for the other endpoint leads to the following:

$$\phi\left(t_0\right) = \phi\left(t_f\right) = 0 \tag{4.78}$$

which are our boundary conditions for the directions $\phi \in C^1\left[t_0, t_f\right]$ of (4.76).

Thus, based on (4.78) and our discussion of unconstrained minimization in infinite-dimensional vector spaces, the appropriate necessary conditions are

$$\delta J\left(x, \phi\right) = \int_{t_0}^{t_f} \left[\frac{\partial f_0\left(x, \dot{x}, t\right)}{\partial x}\phi + \frac{\partial f_0\left(x, \dot{x}, t\right)}{\partial \dot{x}}\dot{\phi}\right] dt = 0 \tag{4.79}$$

$$\phi\left(t_0\right) = \phi\left(t_f\right) = 0 \tag{4.80}$$

for all $\phi \in C^1\left[t_0, t_f\right]$. We may now invoke the fundamental lemma of the calculus of variations presented in Chapter 3; doing show establishes that (4.79) and (4.80) can only hold if

$$\frac{\partial f_0\left(x, \dot{x}, t\right)}{\partial x} - \frac{d}{dt}\left[\frac{\partial f_0\left(x, \dot{x}, t\right)}{\partial \dot{x}}\right] = 0 \tag{4.81}$$

$$x\left(t_0\right) = x_0 \tag{4.82}$$

$$x\left(t_f\right) = x_f \tag{4.83}$$

where (4.81) is recognized as the Euler-Lagrange equation and (4.82) and (4.83) are the original boundary conditions of the problem we have analyzed. Note in particular that these relationships constitute necessary conditions for the calculus of variations problem (4.72), (4.73), and (4.74); they are identical to those we derived in Chapter 3.

4.6 Kuhn-Tucker Conditions for Hilbert Spaces

We have been able so far to uncover several properties of infinite-dimensional mathematical programs that may be viewed as generalizations of results we know to be true for finite-dimensional mathematical programs. It is therefore no surprise that infinite-dimensional programs have necessary conditions that are recognizable generalizations of the finite-dimensional Kuhn-Tucker conditions when a suitable constraint qualification and other regularity conditions are enforced.

Let V be a real Hilbert space, arbitrary elements of which are denoted by v and ϕ, while V^* denotes the dual space to V. Suppose that $J(v)$ is a functional over

V, which is differentiable in the sense of Gâteaux. Furthermore, let the $g_i(v)$ for $i \in [1,m]$ be functionals which are also differentiable in the sense of Gâteaux. We denote the G-derivatives of $J(v)$ and of the $g_i(v)$ by

$$\delta J\ (v,\phi) \quad \text{and} \quad \delta g_i(v,\phi) \quad i \in [1,m] \tag{4.84}$$

for directions $\phi \in V$. These derivatives are considered elements of the dual space V^*. Additionally, we shall assume the G-derivatives (4.84) are continuous linear forms in ϕ, so that

$$\delta J\ (v,\phi) = \langle \nabla J(v), \phi \rangle \tag{4.85}$$

$$\delta g_i(v,\phi) = \langle \nabla g_i(v), \phi \rangle \quad i \in [1,m] \tag{4.86}$$

and

$$\delta J\ (v,\kappa\phi) = \kappa \delta J\ (v,\phi) \tag{4.87}$$

for any scalar κ. The problem we wish to address is of course

$$\min J(v) \quad s.t. \quad v \in U \equiv \{v : g_i(v) \leq 0\ \ i \in [1,m]\} \subset V \tag{4.88}$$

We will invoke a constraint qualification reminiscent of that employed by Kuhn and Tucker (1951) for finite-dimensional mathematical programs and applied by Ritter (1967) to infinite-dimensional programs. That constraint qualification, which serves to exclude certain singularities which might otherwise occur on the boundary of U, is the subject of the following definition:

Definition 4.39. *Kuhn-Tucker constraint qualification for infinite-dimensional mathematical program (4.88). We will say the Kuhn-Tucker constraint qualification holds if there exists a differential mapping $h(t) : [0,1] \longrightarrow V$ with the properties:*

1. $h(t) \in U$ *for* $t \in [0,\alpha] \subset \Re^1_+$;
2. $h(0) = v^*$; *and*
3. $\dfrac{dh\,(0)}{dt} = \beta\phi$ *for* $\beta \in \Re^1_{++}$

for any point $v^* \in U$ *and* $\phi \in V$ *such that* $\delta g_i(v^*,\phi) < 0$ *for all* $i \in I = \{i : g_i(v^*) = 0\}$.

Armed with the foregoing constraint qualification, we are ready to state and prove the following theorem:

Theorem 4.18. *Kuhn-Tucker optimality conditions for mathematical program (4.88). Suppose the constraint qualification expressed in Definition 4.39 holds. Then:*

(i) If $J(v)$ and the $g_i(v)$ for $i \in [1,m]$ are G-differentiable and their G-derivatives are continuous linear forms on the Hilbert space V, then there exist scalar

multipliers η_i for all $i \in [1, m]$ such that the following conditions are necessary for v^* to be a local minimum of (4.88):

$$\nabla J(v^*) + \sum_{i=1}^{m} \eta_i \nabla g_i(v^*) = 0 \tag{4.89}$$

$$\eta_i g_i(v^*) = 0 \quad i \in [1, m] \tag{4.90}$$

$$\eta_i \geq 0 \qquad i \in [1, m] \tag{4.91}$$

(ii) If, in addition, $J(v)$ is convex on U and the $g_i(v)$ are convex on U for all $i \in [1, m]$, then conditions (4.89), (4.90), and (4.91) are also sufficient for v^* to be a global minimum of $J(v)$ on U.

Proof. (i) Let v^* be a local minimum. If $g_i(v^*) < 0$ for each i, then $\delta J(v^*, \phi) = 0$ for all $\phi \in V$; it follows easily that $\nabla J(v^*) = 0$ on U. Thus, for such a circumstance, conditions (4.89), (4.90), and (4.91) are satisfied trivially with

$$\eta_1 = \eta_2 = \cdots = \eta_m = 0$$

Now suppose

$$g_i(v^*) = 0 \quad i \in I$$
$$g_i(v^*) < 0 \quad i \notin I$$

Let ϕ be an arbitrary element of V such that $\delta g_i(v^*, \phi) > 0$ for all $i \in I$ and $\phi \in U$. For the differential mapping h, we know that

$$J(h) = J(v^*) + \delta J(v^*, \frac{dh\,(0)}{dt} t) + O(|\ t\ |) \tag{4.92}$$

By the given, $v = h(t)$ is feasible for $t \in [0, 1]$. If α is a suitably small positive scalar and we set $t = \alpha$ in (4.92), terms of order $o(\alpha)$ are negligible. So we have

$$0 \leq J(h) - J(v^*) = \alpha\beta[\delta J(v^*, \phi)]$$

since $\dot{h}(0) = \beta\phi$ and $\delta J(v^*, \phi)$ is a continuous linear form. Because $\beta > 0$, it follows that

$$\delta J(v^*, \phi) \geq 0 \tag{4.93}$$

Furthermore, by virtue of the constraint qualification, we have

$$\delta g_i(v^*, \phi) > 0 \tag{4.94}$$

So results (4.93) and (4.94) characterizing optimality in the presence of the constraint qualification will hold, if the following system has no solution,:

$$\delta g_i(v^*, \phi) \leq 0 \tag{4.95}$$

$$\delta J(v^*, \phi) < 0 \tag{4.96}$$

Because the G-derivatives are continuous linear forms, we have

$$\delta g_i(v^*, \phi) = \langle \nabla g_i(v^*), \phi \rangle \leq 0 \tag{4.97}$$

$$\delta J(v^*, \phi) = \langle \nabla J(v^*), \phi \rangle < 0 \tag{4.98}$$

By Farkas' lemma, since the system comprised of (4.97) and (4.98) has no solution, it must be that

$$\left[\nabla g(v^*)\right]^T \eta = -\nabla J(v^*) \tag{4.99}$$

$$\eta \geq 0 \tag{4.100}$$

has a solution, where

$$\eta = (\eta_i : i \in I)$$

By defining

$$\eta_i = 0 \quad \forall i \notin I$$

we assure that

$$\nabla J(v^*) + \sum_{i=1}^{m} \eta_i \nabla g_i(v^*) = 0 \tag{4.101}$$

$$\eta_i g_i(v^*) \geq 0 \tag{4.102}$$

$$\eta_i \geq 0 \tag{4.103}$$

(ii) Suppose $(v^*, \eta_1, \ldots, \eta_m)$ satisfies the conditions (4.89), (4.90), and (4.91). Since $J(v)$ is convex by the given

$$J(v^*) + t[J(v) - J(v^*)] \leqq J[v^* + t(v - v^*)] \tag{4.104}$$

holds for any pair (v, v^*) and $t \in [0, 1]$. For $t \in (0, 1]$ the relation (4.104) is equivalent to

$$J(v) - J(v^*) \leqq \frac{1}{t}\{J[v^* - t(v - v^*)] + J(v^*)\}$$

Since $J(v)$ is G-differentiable, it follows that

$$J(v) - J(v^*) \leqq \delta J(v^*, v - v^*) = \langle \nabla J(v^*), v - v^* \rangle \tag{4.105}$$

Similarly, we obtain for each $i \in [1, m]$

$$g_i(v) - g_i(v^*) \geqq \delta g_i(v^*, v - v^*) = \langle \nabla g_i(v^*), v - v^* \rangle \tag{4.106}$$

Because of (4.90), if $\eta_i > 0$, then $g_i(v^*) = 0$; in that case (4.106) yields

$$\langle \nabla g_i(v^*), v - v^* \rangle \leqq g_i(v)$$

Therefore, we see that by exploiting the Kuhn-Tucker identity (4.89)

$$J(v) - J(v^*) \leqq \langle \nabla J(v^*), v - v^* \rangle = \sum_{i=1}^{m} \eta_i \langle \nabla g_i(v^*), v - v^* \rangle \leqq 0$$

for any $v \in U$. This completes the proof. ■

4.7 Mathematical Programming Algorithms

For almost every algorithm for finite-dimensional mathematical programs, there is an analogous algorithm for infinite-dimensional mathematical programs. In this section we study three categories of algorithms for infinite-dimensional mathematical programming:

1. steepest descent methods
2. projected gradient methods
3. penalty function methods

We give detailed statements of continuous-time algorithms belonging to each of these categories; proofs of convergence are also provided. All algorithms are illustrated by application to example problems. The reader should note that, for each algorithm considered, our theoretical presentation and our examples are limited to fixed step sizes.

4.7.1 The Steepest Descent Algorithm

The notion of steepest descent is an algorithmic philosophy for unconstrained optimization wherein we follow the negative gradient of the criterion when minimizing. The algorithm can be applied to infinite-dimensional mathematical programs of the form

$$\min J(u) \quad \text{s.t. } u \in V$$

where V is a Hilbert space. If we take $V = \left(L^2\left[t_0, t_f\right]\right)^m$, the method is applicable to optimal control problems of the form

$$\min J(u) = K\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} f_0(x, u, t)\, dt \tag{4.107}$$

subject to

$$\frac{dx}{dt} = f\left(x, u, t\right) \tag{4.108}$$

$$x\left(t_0\right) = x_0 \tag{4.109}$$

where x_0 is a known, fixed vector and both t_0 and t_f are fixed. Note that

$$H\left(x, u, \lambda, t\right) = f_0(x, u, t) + \lambda^T f\left(x, u, t\right)$$

is the Hamiltonian for the unconstrained problem (4.107), (4.108), and (4.109).

4.7.1.1 Structure of the Steepest Descent Algorithm

The specific algorithmic structure we are considering is:

Steepest Descent Algorithm

Step 0. Initialization. Set $k = 0$ and pick $u^0\left(t\right) \in \left(L^2\left[t_0, t_f\right]\right)^m$.

Step 1. Find state trajectory. Using $u^k\left(t\right)$ solve the state initial-value problem

$$\frac{dx}{dt} = f\left(x, u^0, t\right)$$

$$x\left(t_0\right) = x_0$$

and call the solution $x^k\left(t\right)$.

Step 2. Find adjoint trajectory. Using $u^k\left(t\right)$ and $x^k\left(t\right)$ solve the adjoint final-value problem

$$(-1)\frac{d\lambda}{dt} = \frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial x}$$

$$\lambda\left(t_f\right) = \frac{\partial K\left[x\left(t_f\right), t_f\right]}{\partial x}$$

and call the solution $\lambda^k\left(t\right)$.

Step 3. Find gradient. Using $u^k\left(t\right)$, $x^k\left(t\right)$ and $\lambda^k\left(t\right)$ calculate

$$\nabla_u J(u^k) = \left[\frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial u}\right]^T$$

$$\frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial u} = \frac{\partial f_0(x^k, u^k, t)}{\partial u} + \left(\lambda^k\right)^T \frac{\partial f\left(x^k, u^k, t\right)}{\partial u}$$

Step 4. Update and apply stopping test. For a suitably small step size θ_k, update according to

$$u^{k+1} = u^k - \theta_k \nabla_u J(u^k)$$

If an appropriate stopping test is satisfied, declare

$$u^* (t) \approx u^{k+1} (t)$$

Otherwise set $k = k + 1$ and go to Step 1.

4.7.1.2 Convergence of the Steepest Descent Algorithm

To establish convergence of the steepest descent algorithm, it is helpful to first establish two lemmas. The first preliminary result is the following:

Lemma 4.1. *Suppose the functional $J : V \longrightarrow \Re^1$ has a well-defined gradient on V, a reflexive Banach space. Take $\nabla J\,(u)$ to be uniformly continuous, and define*

$$g_k(\theta) \equiv J\left(u^k + \theta d^k\right)$$

Let the step size θ_k be defined by

$$\begin{aligned} \theta_k &> 0 \\ g_k'(\theta_k) &= 0 \\ J\left(u^k + \theta_k d^k\right) &\leq J\left(u^k + \theta d^k\right) \quad \forall \theta \in [0, \theta_k] \end{aligned}$$

Then

$$\theta_k > \eta\left(\left|c \cdot g_k'\,(0)\right|\right) \quad \forall c \in (0,1) \tag{4.110}$$

where $\eta\left(\left|c \cdot g_k'\,(0)\right|\right)$ satisfies

$$\left\|u^{k+1} - u^k\right\| \leq \eta\left(\left|c \cdot g_k'\,(0)\right|\right)$$

Proof. The assumption of uniform continuity means

$$\|u - v\| \leq \eta\,(\varepsilon) \Longrightarrow |\langle \nabla J\,(u) - \nabla J\,(v)\,, \phi\rangle| \leq \varepsilon \quad \forall \phi \text{ such that } \|\phi\| = 1 \tag{4.111}$$

while

$$\eta\,(\varepsilon) \longrightarrow 0_+ \Longleftrightarrow \varepsilon \longrightarrow 0_+ \tag{4.112}$$

Following Minoux (1986) we define

$$g_k\,(\theta) = J\left(u^k + \theta d^k\right) \tag{4.113}$$

Then we have

$$g_k'(\theta) = \frac{d}{d\theta} J\left(u^k + \theta d^k\right) = \left\langle \nabla J\left(u^k + \theta d^k\right), d^k \right\rangle \tag{4.114}$$

We assert that

$$\theta_k > \frac{\eta\left(\left|c \cdot g_k'(0)\right|\right)}{\left\|d^k\right\|} = \eta\left(\left|c \cdot g_k'(0)\right|\right) \quad \forall c \in (0,1) \tag{4.115}$$

when we assume $\left\|d^k\right\| = 1$. To prove this assertion, suppose

$$\theta_k \leq \eta\left(\left|c \cdot g_k'(0)\right|\right) \tag{4.116}$$

Then, note

$$u^{k+1} - u^k = \theta_k d^k \tag{4.117}$$

From (4.116) and (4.117), we have

$$\left\|u^{k+1} - u^k\right\| = \theta_k \left\|d^k\right\| \leq \eta\left(\left|c \cdot g_k'(0)\right|\right)$$

If we take $\varepsilon = \left|c \cdot g_k'(0)\right|$, the precondition of (4.111) is met. Therefore, also by (4.111), we have

$$\begin{aligned} \left|c \cdot g_k'(0)\right| &\geq \left|\left\langle \nabla J\left(u^k + \theta_k d^k\right) - \nabla J\left(u^k\right), d^k \right\rangle\right| \\ &= \left|\left\langle \nabla J\left(u^k + \theta_k d^k\right), d^k \right\rangle - \left\langle \nabla J\left(u^k\right), d^k \right\rangle\right| \\ &= \left|g_k'(\theta_k) - g_k'(0)\right| = \left|g_k'(0)\right| \end{aligned} \tag{4.118}$$

That is

$$\left|c \cdot g_k'(0)\right| \geq \left|g_k'(0)\right| \tag{4.119}$$

which is a contradiction since $c \in (0,1)$. Thus, assertion (4.115) holds. ■

The following result, which depends on Lemma 4.1, will be directly employed in the convergence proof:

Lemma 4.2. *For the given of Lemma 4.1, the following inequality obtains*

$$\Delta J_k\left(u^k; \theta_k\right) \geq \Delta J_k\left(u^k; \bar{\theta}_k\right) \geq \eta\left(\left|c \cdot g_k'(0)\right|\right)(1-c)\left|g_k'(0)\right| \tag{4.120}$$

where

$$\begin{aligned} \bar{\theta}_k &\equiv \eta\left(\left|c \cdot g_k'(0)\right|\right) \\ \Delta J_k\left(u^k; \theta\right) &\equiv J\left(u^k\right) - J\left(u^k + \theta d^k\right) \quad \forall \theta \in [0, \theta_k] \end{aligned}$$

Proof. By the given, we know (4.115) obtains. Also by virtue of the given, we have the following condition for the step size

$$J\left(u^k+\theta_k d^k\right)\leq J\left(u^k+\theta d^k\right)\quad \forall\theta\in[0,\theta_k]$$

from which it follows that

$$\Delta J_k\left(u^k;\theta\right)=J\left(u^k\right)-J\left(u^k+\theta d^k\right) \tag{4.121}$$

$$\leq J\left(u^k\right)-J\left(u^k+\theta_k d^k\right)=\Delta J_k\left(u^k;\theta_k\right) \tag{4.122}$$

That is

$$\Delta J_k\left(u^k;\theta\right)\leq\Delta J_k\left(u^k;\theta_k\right)\quad \forall\theta\in[0,\theta_k] \tag{4.123}$$

In (4.123) choose

$$\theta=\bar{\theta}_k\equiv\eta\left(\left|c\cdot g_k'(0)\right|\right) \tag{4.124}$$

so that from (4.115) we have

$$\bar{\theta}_k<\theta_k \tag{4.125}$$

Then from (4.123) and (4.125) we have

$$\Delta J_k\left(u^k;\bar{\theta}_k\right)\leq\Delta J_k\left(u^k;\theta_k\right) \tag{4.126}$$

By virtue of (4.113) and (4.122), we have

$$\Delta J_k\left(u^k;\bar{\theta}_k\right)=J\left(u^k\right)-J\left(u^k+\bar{\theta}_k d^k\right)=g_k(0)-g_k\left(\bar{\theta}_k\right) \tag{4.127}$$

By the mean-value theorem, we know

$$g_k(0)-g_k\left(\bar{\theta}_k\right)=\bar{\theta}_k g_k'\left(\theta_k^0\right)\quad \text{for some}\ \ \theta_k^0\in\left(0,\bar{\theta}_k\right) \tag{4.128}$$

Substitution of (4.128) into (4.127) yields

$$\Delta J_k\left(u^k;\bar{\theta}_k\right)=\bar{\theta}_k g_k'\left(\theta_k^0\right) \tag{4.129}$$

Moreover, we have

$$\theta_k^0<\bar{\theta}_k=\eta\left(\left|c\cdot g_k'(0)\right|\right) \tag{4.130}$$

Let us return to (4.111) and set

$$\begin{aligned}u&=u^k+\theta_k^0\\ v&=u^k\\ \varepsilon&=\left|c\cdot g_k'(0)\right|\\ \phi&=d^k\end{aligned}$$

so that

$$\left\| u^k + \theta_k^0 - u^k \right\| = \theta_k^0 \leq \eta \left(\left| c \cdot g_k' (0) \right| \right) \tag{4.131}$$

$$\Longrightarrow \left| \left\langle \nabla J \left(u^k + \theta_k^0 \right) - \nabla J \left(u^k \right), d^k \right\rangle \right| \leq \left| c \cdot g_k' (0) \right| \tag{4.132}$$

Because of (4.130), we know (4.131) holds; therefore (4.132) also holds. Moreover, (4.114) allows us to restate (4.132) as

$$\left| g_k' \left(\theta_k^0 \right) - g_k' (0) \right| \leq \left| c \cdot g_k' (0) \right| \tag{4.133}$$

which in turn yields

$$- \left| c \cdot g_k' (0) \right| \leq g_k' \left(\theta_k^0 \right) - g_k' (0) \leq \left| c \cdot g_k' (0) \right|$$

or

$$g_k' \left(\theta_k^0 \right) \geq g_k' (0) - \left| c \cdot g_k' (0) \right|$$

If $g_k' (0) \geq 0$ then

$$g_k' \left(\theta_k^0 \right) \geq (1 - c) \, g_k' (0) \tag{4.134}$$

If $g_k' (0) < 0$ then

$$g_k' \left(\theta_k^0 \right) \geq (1 + c) \, g_k' (0) \tag{4.135}$$

Since $0 < c < 1$, the inequality

$$g_k' \left(\theta_k^0 \right) \geq (1 - c) \left| g_k' (0) \right| \tag{4.136}$$

ensures both case (4.134) and case (4.135) are satisfied. Now we note that, taken together, (4.129) and (4.136) give

$$\Delta J_k \left(u^k ; \bar{\theta}_k \right) = \bar{\theta}_k g_k' \left(\theta_k^0 \right) \geq \bar{\theta}_k \, (1 - c) \left| g_k' (0) \right| \tag{4.137}$$

In (4.124), we set $\bar{\theta}_k = \eta \left(\left| c \cdot g_k' (0) \right| \right)$, so that the following inequality

$$\Delta J_k \left(u^k ; \theta_k \right) \geq \Delta J_k \left(u^k ; \bar{\theta}_k \right) \geq \eta \left(\left| c \cdot g_k' (0) \right| \right) (1 - c) \left| g_k' (0) \right| \tag{4.138}$$

is immediate from (4.137). ■

We are now ready to present the main convergence result for the steepest descent algorithm:

Theorem 4.19. *Convergence of the steepest descent algorithm. Suppose the functional $J : V \longrightarrow \Re^1$ has a well-defined gradient and is convex and weakly bounded*

from below, where V is a reflexive Banach space. Take $\nabla J(u)$ to be uniformly continuous, and determine the optimal step size θ_k according to

$$1 > \theta_k > 0 \tag{4.139}$$

$$d^k = -\frac{\nabla J\left(u^k\right)}{\left\|\nabla J\left(u^k\right)\right\|} \tag{4.140}$$

$$\frac{d}{d\theta} J\left(u^k + \theta_k d^k\right) = \langle \nabla J\left(u^k + \theta_k d^k\right), d^k \rangle = 0 \tag{4.141}$$

$$J\left(u^k + \theta_k d^k\right) \leq J\left(u^k + \theta d^k\right) \quad \forall \theta \in [0, \theta_k] \tag{4.142}$$

Then, if the condition

$$\lim_{\|u\| \longrightarrow \infty} J(u) \longrightarrow \infty \tag{4.143}$$

holds, the steepest descent algorithm converges to a minimum u^* of J on V.

Proof. Following Minoux (1986), we again recall from the given that

$$J\left(u^k + \theta_k d^k\right) \leq J\left(u^k + \theta d^k\right) \quad \forall \theta \in [0, \theta_k]$$

Choosing $\theta = 0$, we obtain

$$J\left(u^k + \theta_k d^k\right) \leq J\left(u^k\right)$$

which we restate as

$$J\left(u^{k+1}\right) \leq J\left(u^k\right)$$

Hence $\left\{J\left(u^k\right)\right\}$ is a decreasing sequence. Thus, by the assumption that J is bounded from below, we have

$$\lim_{k \longrightarrow \infty} \left\{J\left(u^k\right) - J\left(u^{k+1}\right)\right\} = 0$$

This condition leads to

$$\begin{aligned} \lim_{k \longrightarrow \infty} \left\{J\left(u^k\right) - J\left(u^{k+1}\right)\right\} &= \lim_{k \longrightarrow \infty} \left\{J\left(u^k\right) - J\left(u^k + \theta_k d^k\right)\right\} \\ &= \lim_{k \longrightarrow \infty} \Delta J_k\left(u^k; \theta_k\right) = 0 \end{aligned} \tag{4.144}$$

As $\Delta J_k\left(u^k; \theta_k\right)$ approaches zero, we see from Lemma 4.2, specifically from (4.138), that

$$\lim_{k \longrightarrow \infty} \left|g_k'(0)\right| \eta\left(\left|c \cdot g_k'(0)\right|\right) = 0 \tag{4.145}$$

which in turn requires

$$\lim_{k \longrightarrow \infty} \left|g_k'(0)\right| = 0 \tag{4.146}$$

By (4.114) of the proof of Lemma 4.1, we know that

$$g_k'(0) = \langle J\left(u^k\right), d^k\rangle \tag{4.147}$$

From (4.146) and (4.147), it is immediate that

$$\lim_{k\longrightarrow\infty}\left|\left\langle \nabla J\left(u^k\right), d^k\right\rangle\right| = 0 \tag{4.148}$$

Thus, we have

$$\begin{aligned}\lim_{k\longrightarrow\infty}\left|\left\langle \nabla J\left(u^k\right), d^k\right\rangle\right| &= \lim_{k\longrightarrow\infty}\left|\left\langle \nabla J\left(u^k\right), -\frac{\nabla J\left(u^k\right)}{\left\|\nabla J\left(u^k\right)\right\|}\right\rangle\right| \\ &= \lim_{k\longrightarrow\infty}\left|-\frac{\left\|\nabla J\left(u^k\right)\right\|^2}{\left\|\nabla J\left(u^k\right)\right\|}\right| \\ &= \lim_{k\longrightarrow\infty}\left\|\nabla J\left(u^k\right)\right\| = 0,\end{aligned} \tag{4.149}$$

Furthermore, the assumption $\lim_{\|u\|\longrightarrow\infty} J(u) \longrightarrow \infty$, together with the already established fact that the sequence $\left\{J(u^k)\right\}$ is decreasing, implies that all u^k are contained in a bounded set. Our assumption that $J(u)$ is weakly bounded, allows us to apply the weak compactness theorem (Theorem 4.3) to conclude that there exists at least one weak cluster point, u^*. That is, there exists a sequence $u^k \longrightarrow u^*$ (weakly) as $k \longrightarrow \infty$ for some $u^* \in V$. From the convexity assumption, we have

$$J(v) \geq J\left(u^k\right) + \left\langle \nabla J\left(u^k\right), v - u^k\right\rangle \qquad \forall v \in V \tag{4.150}$$

Note further that

$$\lim_{k\longrightarrow\infty}\left\langle \nabla J\left(u^k\right), v - u^k\right\rangle \longrightarrow 0 \tag{4.151}$$

since $\left\|\nabla J\left(u^k\right)\right\| \longrightarrow 0$ and u^k is contained in a bounded set while: (1) either $v = u^\ell \in \left\{u^k\right\}$ and is therefore bounded or (2) $v \notin \left\{u^k\right\}$ and is therefore unaffected by the limit $k \longrightarrow \infty$. From (4.150) and (4.151) it is immediate that

$$J(v) \geq \lim_{k\longrightarrow\infty} J\left(u^k\right) \qquad \forall v \in V \tag{4.152}$$

Moreover, since J is convex, J is weakly lower semicontinuous; therefore, by Definition 4.17, we have

$$u^k \longrightarrow u^* \text{ (weakly)} \Longrightarrow \lim_{k\longrightarrow\infty} J\left(u^k\right) \geq J\left(u^*\right) \tag{4.153}$$

From (4.152) and (4.153) we have

$$J(v) \geq J\left(u^*\right) \qquad \forall v \in V \tag{4.154}$$

This completes the proof. ■

4.7.2 The Projected Gradient Algorithm

In this section we will be concerned with constructing an algorithm for the constrained infinite-dimensional mathematical program:

$$\min J(u) \quad s.t.\ u \in U \subset V \tag{4.155}$$

where V is a Hilbert space and

$$J : V \longrightarrow \Re_+^1$$

Furthermore, we suppose that $J(u)$ is G-differentiable on U and that $V = \left(L^2\left[t_0, t_f\right]\right)^m$. In this case, V is a Hilbert space, and we know that the G-derivative of the functional $J(u)$ is well defined and allows immediate articulation of a first-order necessary for the optimal solution $u^* \in U$:

$$\delta J\left(u^*, \phi\right) = \left\langle \nabla J\left(u^*\right), u - u^* \right\rangle \geq 0 \quad \forall u \in U \tag{4.156}$$

It therefore seems reasonable to explore algorithms based on the notion of the gradient of a functional. Clearly, if there were no constraints ($U = V$), we could directly employ the steepest descent algorithm. However, many applied problems of interest have constraints, so we need some notion of modifying the gradient direction when it points out of U in order to obtain a related, alternative direction that is feasible. This is accomplished by use of the minimum norm projection.

4.7.2.1 The Minimum Norm Projection

Consider the mathematical program

$$\min \left\| u - u^k \right\| \quad s.t.\ u \in U \subset V \tag{4.157}$$

where V is a Hilbert space and $\|v\| = \left\langle v^T, v \right\rangle^{\frac{1}{2}}$ of course denotes the norm induced by the scalar product. We say that

$$z^k = P_U\left[u - u^k\right] = \arg\min_{u \in U}\left(\left\| u - u^k \right\|\right) \tag{4.158}$$

is the *minimum norm projection of* u^k *onto* U. It is equivalent to write

$$z^k = \arg\min_{u \in U} \frac{1}{2} \left\| u - u^k \right\|^2 \tag{4.159}$$

$$= \arg\min_{u \in U} \left[J^k_{\text{proj}}(u) = \frac{1}{2} \left(u - u^k \right)^T \cdot \left(u - u^k \right) \right] \tag{4.160}$$

$$\equiv \arg\min_{u \in U} \int_0^1 \frac{1}{2} \left(u - u^k \right)^2 dt \tag{4.161}$$

since the norm is a monotonic transformation. If $V = \left(L^2 \left[t_0, t_f \right] \right)^m$, we know that the gradient of the objective functional in (4.161) is

$$\nabla J^k_{\text{proj}}(u) = \left(u - u^k \right)$$

Clearly, a necessary condition that z^k must satisfy is

$$\left\langle \nabla J^k_{\text{proj}} \left(z^k \right), \left(u - z^k \right) \right\rangle \geq 0 \quad \forall u \in U$$

which is equivalent to

$$\left\langle \left(z^k - u^k \right), \left(u - z^k \right) \right\rangle \geq 0 \quad \forall u \in U \tag{4.162}$$

Due to the convexity of the minimum norm problem, necessary condition (4.162) is also a sufficient condition. This variational inequality will be a key ingredient in stating and proving the convergence of a projection algorithm for constrained mathematical programming in $\left(L^2 \left[t_0, t_f \right] \right)^m$.

The first thing we want to show is that the minimum norm projection is a contraction mapping; that is, we want to show that for one iteration

$$\left\| z^1 - z^0 \right\| \leq \left\| u^1 - u^0 \right\| \tag{4.163}$$

This result can be proven using the variational inequality (4.162) together with Schwartz's inequality for Hilbert spaces:

$$\|a\| \, \|b\| \geq \langle a, b \rangle \tag{4.164}$$

From (4.162) we can state immediately the following two results for all $u \in U$:

$$\left\langle \left(z^0 - u^0 \right), \left(u - z^0 \right) \right\rangle \geq 0 \tag{4.165}$$

$$\left\langle \left(z^1 - u^1 \right), \left(u - z^1 \right) \right\rangle \geq 0 \tag{4.166}$$

In (4.165) set $u = z^1 \in U$ and in (4.166) set $u = z^0 \in U$, as we are free to do, to obtain

$$\left\langle \left(z^0 - u^0\right), \left(z^1 - z^0\right) \right\rangle \geq 0 \tag{4.167}$$

$$\left\langle \left(z^1 - u^1\right), \left(z^0 - z^1\right) \right\rangle \geq 0 \tag{4.168}$$

Adding these last two expressions yields

$$\left\langle \left(z^0 - u^0\right) + \left(u^1 - z^1\right), \left(z^1 - z^0\right) \right\rangle \geq 0$$

which leads to

$$\left\langle \left(z^0 - z^1\right) + \left(u^1 - u^0\right), \left(z^1 - z^0\right) \right\rangle \geq 0$$

Further manipulation gives

$$\left\langle \left(u^1 - u^0\right), \left(z^1 - z^0\right) \right\rangle - \left\langle \left(z^1 - z^0\right), \left(z^1 - z^0\right) \right\rangle \geq 0$$

which is equivalent to

$$\left\langle \left(u^1 - u^0\right), \left(z^1 - z^0\right) \right\rangle - \left\| z^1 - z^0 \right\|^2 \geq 0 \tag{4.169}$$

By Schwartz's inequality we know that

$$\left\langle \left(u^1 - u^0\right), \left(z^1 - z^0\right) \right\rangle \leq \left\| u^1 - u^0 \right\| \cdot \left\| z^1 - z^0 \right\| \tag{4.170}$$

Combining (4.169) and (4.170) gives us

$$\left\| u^1 - u^0 \right\| \cdot \left\| z^1 - z^0 \right\| \geq \left\| z^1 - z^0 \right\|^2 \tag{4.171}$$

from which (4.163) follows immediately. ■

4.7.2.2 Structure of the Gradient Projection Algorithm

The gradient projection algorithm is summarized by the following updating rule:

$$u^{k+1} = P_U \left[u^k - \theta_k \nabla J \left(u^k \right) \right] \tag{4.172}$$

where the superscript k denotes an iteration index and θ_k is a step size for iteration k. Note that this is fundamentally the same projection algorithm familiar from finite-dimensional mathematical programming, but it is now carried out in a Hilbert space V of which U is a subset; the mechanics of the projection are governed by the variational inequality (4.162).

Based on (4.172) we may now provide the following formal statement of the gradient projection algorithm:

Gradient Projection Algorithm

Step 0. Initialization. Set $k = 0$ and pick $u^0(t) \in \left(L^2\left[t_0, t_f\right]\right)^m$.

Step 1. Find state trajectory. Using $u^k(t)$ solve the state initial-value problem

$$\frac{dx}{dt} = f\left(x, u^0, t\right)$$
$$x(t_0) = x_0$$

and call the solution $x^k(t)$.

Step 2. Find adjoint trajectory. Using $u^k(t)$ and $x^k(t)$ solve the adjoint final value problem

$$(-1)\frac{d\lambda}{dt} = \frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial x}$$
$$\lambda\left(t_f\right) = \frac{\partial K\left[x\left(t_f\right), t_f\right]}{\partial x}$$

and call the solution $\lambda^k(t)$.

Step 3. Find gradient. Using $u^k(t)$, $x^k(t)$, and $\lambda^k(t)$ calculate

$$\nabla_u J(u^k) = \frac{\partial H\left(x^k, u^k, \lambda, t\right)}{\partial u}$$
$$= \frac{\partial f_0(x^k, u^k, t)}{\partial u} + \left(\lambda^k\right)^T \frac{\partial f\left(x^k, u^k, t\right)}{\partial u}$$

Step 4. Update and apply stopping test. For a suitably small step size θ_k, update according to

$$u^{k+1} = P_U\left[u^k - \theta_k \nabla J\left(u^k\right)\right]$$

If an appropriate stopping test is satisfied, declare

$$u^*(t) \approx u^{k+1}(t)$$

Otherwise, set $k = k + 1$ and go to Step 1.

4.7.2.3 Coerciveness

We need to define a concept known as *coerciveness* or α-convexity that will be important to proving the convergence of the projection algorithm described above:

Definition 4.40. *Let $J : V \longrightarrow \Re^1$ be a functional on V, a normed vector space. J is said to be coercive (or α-convex) if there is a real scalar $\alpha > 0$ such that*

$$J\left[(1-\theta)u + \theta v\right] \leq (1-\theta)\, J(u) + \theta J(v) - \frac{\alpha}{2}\theta(1-\theta)\|u-v\|^2 \tag{4.173}$$

for all $u, v \in V$ and $\theta \in (0,1)$.

Note that if $\alpha = 0$ in (4.173) we have the usual notion of convexity. We also state without proof the following lemma.

Lemma 4.3. *If $J : V \longrightarrow \Re^1$ is G-differentiable at every point of V, then*

$$J \text{ coercive } (\alpha\text{-convex}) \Longleftrightarrow \delta J(u, u-v) - \delta J(v, u-v) \geq \alpha\|u-v\|^2 \tag{4.174}$$

for all $u, v \in V$.

Note that, when the gradient of J is defined on V, the righthand side of (4.174) becomes

$$\langle \nabla J(u) - \nabla J(v), u-v\rangle \geq \alpha\|u-v\|^2 \tag{4.175}$$

4.7.2.4 Convergence of the Gradient Projection Algorithm

We now state and prove the following key result regarding convergence of the projection algorithm:

Theorem 4.20. *Suppose the functional $J : U \subset V \longrightarrow \Re^1$ is coercive (α-convex) with $\alpha > 0$ and $\nabla J(u)$ is defined and satisfies the Lipschitz condition*

$$\|\nabla J(u) - \nabla J(v)\| \leq \beta\|u-v\| \tag{4.176}$$

for all $u, v \in U$. Then the projection algorithm based on the negative gradient direction converges to the minimum u^* of J on U for fixed step size choices

$$\theta \in \left(0, \frac{2\alpha}{\beta^2}\right) \tag{4.177}$$

Proof. We begin by invoking the variational inequality first-order condition

$$\left\langle \nabla J\left(u^*\right), u-u^*\right\rangle \geq 0 \quad \forall u \in U \tag{4.178}$$

Because of variational inequality (4.178) we have

$$\left\langle \left[u^* - \left(u^* - \theta \nabla J\left(u^*\right)\right)\right], \left(u - u^*\right)\right\rangle \geq 0 \quad \forall u \in U \tag{4.179}$$

for all $\theta > 0$. Recalling property (4.162), we see from (4.179) that u^* must be the projection of $u^* - \theta \nabla J\left(u^*\right)$; that is

$$u^* = P_U\left[u^* - \theta \nabla J\left(u^*\right)\right] \tag{4.180}$$

From this last observation and the projection algorithm itself (4.172), it is an easy matter to construct the difference

$$u^{k+1} - u^* = P_U\left[u^k - \theta \nabla J\left(u^k\right)\right] - P_U\left[u^* - \theta \nabla J\left(u^*\right)\right] \tag{4.181}$$

Remembering result (4.163) that establishes the projection mapping is a contraction, we obtain from (4.181)

$$\left\|u^{k+1} - u^*\right\| \leq \left\|\left(u^k - \theta \nabla J\left(u^k\right)\right) - \left(u^* - \theta \nabla J\left(u^*\right)\right)\right\| \tag{4.182}$$

so that

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left\|\left(u^k - u^*\right) - \theta\left[\nabla J\left(u^k\right) - \nabla J\left(u^*\right)\right]\right\|^2 \tag{4.183}$$

The righthand side (RHS) of (4.183) can be restated using the given Lipschitz condition and property (4.175):

$$\begin{aligned} RHS &= \left\|u^k - u^*\right\|^2 - 2\theta\left\langle \nabla J\left(u^k\right) - \nabla J\left(u^*\right), u^k - u^*\right\rangle \\ &\quad + \theta^2\left\|\nabla J\left(u^k\right) - \nabla J\left(u^*\right)\right\|^2 \\ &\leq \left\|u^k - u^*\right\|^2 - 2\alpha\theta\left\|u^k - u^*\right\|^2 + \beta^2\theta^2\left\|u^k - u^*\right\|^2 \\ &= \left(1 - 2\alpha\theta + \beta^2\theta^2\right)\left\|u^k - u^*\right\|^2 \end{aligned} \tag{4.184}$$

Results (4.183) and (4.184) tell us that

$$\begin{aligned} \left\|u^{k+1} - u^*\right\|^2 &\leq \left(1 - 2\alpha\theta + \beta^2\theta^2\right)\left\|u^k - u^*\right\|^2 \\ &\Longrightarrow \left\|u^{k+1} - u^*\right\| \leq \left(1 - 2\alpha\theta + \beta^2\theta^2\right)^{\frac{1}{2}}\left\|u^k - u^*\right\| \end{aligned} \tag{4.185}$$

Inequality (4.185) will establish convergence if

$$\begin{aligned} \left(1 - 2\alpha\theta + \beta^2\theta^2\right) &< 1 \\ &\Longrightarrow \beta^2\theta^2 < 2\alpha\theta \\ &\Longrightarrow \theta < \frac{2\alpha}{\beta^2} \end{aligned} \tag{4.186}$$

Since $\theta > 0$, the desired result follows. ■

4.7.3 Penalty Function Methods

When constraints that destroy special structure or other desirable properties of infinite-dimensional mathematical programs arise, one approach is to form penalty functions for these constraints and append them to the objective functional. It is in principle possible to convert a constrained optimization problem into a sequence of unconstrained problems whose solutions converge to the solution of the original problem. As penalty functions are defined below, algorithms based on them begin with infeasible points and move toward feasibility, although feasibility is generally achieved only in the sense of a limit.

4.7.3.1 Definition of a Penalty Function

Let us first recall what is meant by weak lower semicontinuity according to Definition 4.17: the functional J is weakly lower semicontinuous on V if for all $v, v^k \in V$ such that

$$v^k \rightarrow v \text{ (weakly)},$$

we have

$$\lim_{k \longrightarrow \infty} \inf J\left(v^k\right) \geq J(v)$$

Now consider

$$\min J(u) \quad s.t.\, u \in U \subset V \tag{4.187}$$

Furthermore, we say that $P(u)$ is a penalty function for (4.187) if

$$\begin{aligned} &P(u) = 0 \iff u \in U \\ &P(u) \geq 0 \quad \forall u \\ &P \text{ is weakly lower semicontinuous} \end{aligned}$$

Our intention is to replace the constrained problem (4.187) with the unconstrained problem

$$\min\ J_\rho(u) = J(u) + \rho P(u) \tag{4.188}$$

where $\rho > 0$ is a penalty multiplier (parameter) that tends to infinity so that the product $\rho P(u)$ is recognized as positive according to the numerical precision of the

computing platform employed as the boundary is approached. The following are two examples of penalty functions for minimization problems:

Constraint	Penalty function
$g(x) \leq 0$	$P(x) = \frac{1}{2}[\max(g(x), 0)]^2$
$h(x) = 0$	$P(x) = \frac{1}{2}[h(x)]^2$

4.7.3.2 Description of the Penalty Function Algorithm

In the penalty function method, where the superscript k is an iteration index, we follow the scheme

$$u^k = \arg\left\{\min\ J_{\rho_k}(u) \ = J(u) + \rho_k P(u)\right\} \tag{4.189}$$

given the penalty function multiplier ρ_k, which must be made increasingly large. If an appropriate stopping test is satisfied at iteration k, declare

$$u^*(t) \approx u^k(t)$$

Otherwise, select

$$\rho_{k+1} > \rho_k,$$

set $k = k + 1$, and repeat (4.189).

4.7.3.3 Convergence of the Penalty Function Method

Under certain conditions, (4.188) has a solution u_ρ^* for each $\rho > 0$. In that case, it is our hope that, when $\rho \to \infty$, the sequence $\{u_\rho^*\}$ converges to a solution of (4.187), namely u^*. In particular, we have the following result:

Theorem 4.21. *Assume that J is weakly lower semicontinous, bounded from below, and $J(u) \to \infty$ as $\|u\| \to \infty$. Also assume that the set U is weakly closed. Then every weak cluster point of the sequence $\{u_\rho^*\}$ is an optimal solution of (4.187).*

Proof. We follow Minoux (1986). Note that

$$J_\rho(u_\rho^*) = J\left(u_\rho^*\right) + \rho P(u_\rho^*) \leq J(u_0) + \rho P(u_0) \qquad \forall u_0 \in V$$

Since $P(u_0) = 0$ for $u_0 \in U$, we have

$$J_\rho(u_\rho^*) \leq J(u_0) \quad \forall u_0 \in U$$

Consequently J_ρ remains bounded from above by a constant scalar independent of ρ. Since $J(u) \to \infty$ as $\|u\| \to \infty$, it follows that the whole sequence $\{u_\rho^*\}$ is contained in a bounded set. Using the weak compactness theorem, one may extract from this sequence a subsequence $\{u_{\rho'}^*\}$ which converges weakly to $u^* \in V$. Also, since we have

$$J(u_{\rho'}^*) + \rho' P(u_{\rho'}^*) \leq J(u_0)$$

for all ρ', we may state that

$$P(u_{\rho'}^*) \leq \frac{1}{\rho'} \left[J(u_0) - J(u_{\rho'}^*) \right]$$

Because J is bounded from below, there has to exist a number m_0 such that

$$J(u) \geq m_0 \quad \forall u$$

Consequently

$$P(u_{\rho'}^*) \leq \frac{1}{\rho'} \left[J(u_0) - m_0 \right] \quad \forall \rho'$$

Therefore $P(u_{\rho'}^*) \to 0$ when $\rho' \to \infty$ as required. Because P is weakly lower semicontinous and since $u_{\rho'}^*$ weakly converges to u^*, we know

$$\lim_{\rho' \to \infty} P(u_{\rho'}^*) \geq P\left(u^*\right)$$

Hence $P(u^*) \leq 0$. By the definition of a penalty function, it must also be that $P(u^*) \geq 0$. It follows immediately that $P(u^*) = 0$ and $u^* \in U$. Finally, since

$$J(u_{\rho'}^*) \leq J(u_0) \quad \forall u_0 \in U,$$

it must be that J is weakly lower semicontinuous:

$$J(u^*) \leq \lim_{\rho' \to \infty} \inf J(u_{\rho'}^*) \leq J\left(u_0\right) \quad \forall u_0 \in U$$

It of course then follows that u^* is optimal to problem (4.187). ■

4.7.4 Example of the Steepest Descent Algorithm

Consider the following familiar problem where $u \in L^2[a, b]$ and $x \in \mathcal{H}^1[a, b]$:

$$\min J = \int_a^b \frac{1}{2} \left(x^2 + u^2\right) dt \tag{4.190}$$

subject to

$$\frac{dx}{dt} = Bu \quad (\lambda) \tag{4.191}$$

$$x(a) = A \tag{4.192}$$

We know

$$H = \frac{1}{2}\left(x^2 + u^2\right) + \lambda Bu$$

$$-\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = x$$

$$\lambda(b) = 0$$

We employ the parameter values

$$A = 1.5431$$

$$B = 1$$

$$a = 0$$

$$b = 1$$

To apply the steepest descent algorithm, we must compute the gradient and take steps along it. We note that the gradient of the criterion functional is

$$\nabla_u J(u) = \lambda \frac{\partial f}{\partial u} + \frac{\partial f_0}{\partial u} \tag{4.193}$$

where f refers to the right hand side of the state dynamics and f_0 is the integrand of the criterion. Thus, for the problem at hand

$$\frac{\partial f}{\partial u} = \frac{\partial u}{\partial u} = 1 \tag{4.194}$$

$$\frac{\partial f_0}{\partial u} = \frac{1}{2}\frac{\partial\left[x^2 + u^2\right]}{\partial u} = u \tag{4.195}$$

$$\Longrightarrow \nabla_u J(u) = \lambda + u \tag{4.196}$$

We also know

$$\frac{dx}{dt} = u \tag{4.197}$$

$$\frac{d\lambda}{dt} = -x \tag{4.198}$$

So the iterative procedure that is the steepest descent algorithm in continuous time has the following structure:

1. guess $u^0(t)$
2. calculate $x^0(t)$ using $u^0(t)$ in (4.197),

3. calculate $\lambda^0(t)$ using $x^0(t)$ in (4.198)
4. calculate $\nabla_u J\left(u^0\right) = \lambda^0 + u^0$
5. apply the steepest descent algorithm for step size θ_0:

$$u^1(t) = u^0(t) - \theta_0 \nabla_u J\left[u^0(t)\right] \tag{4.199}$$
$$= u^0(t) - \theta_0\left[\lambda^0(t) + u^0(t)\right] \tag{4.200}$$

6. repeat for subsequent iterations

The following calculations are grouped by iteration number k:

$\boxed{k=0}$: We select $u^0 = 0$; hence $x^0(t) = x(0) = 1.5431$. Therefore

$$\begin{aligned} \dot{\lambda} &= -1.5431 \\ \lambda(1) &= 0 \end{aligned} \tag{4.201}$$

whose solution is

$$\lambda^0(t) = -1.5431t + 1.5431$$

Consequently

$$\nabla_u J\left(u^0\right) = -1.5431t + 1.5431 \tag{4.202}$$
$$u^1(t) = u^0(t) - \theta_0\left[\nabla_u J\left(u^0\right)\right] \tag{4.203}$$
$$= 1.5431\theta_0 t - 1.5431\theta_0 \tag{4.204}$$

$\boxed{k=1}$:

$$\begin{aligned} \dot{x} = u^1 &= 1.5431\theta_0 t - 1.5431\theta_0 \\ x(0) &= 1 \end{aligned} \tag{4.205}$$

whose solution is

$$x^1(t) = .77155\theta_0 t^2 - 1.5431\theta_0 t + 1 \tag{4.206}$$

Consequently

$$\begin{aligned} \dot{\lambda} &= -.77155\theta_0 t^2 + 1.5431\theta_0 t - 1 \\ \lambda(1) &= 0 \end{aligned} \tag{4.207}$$

whose solution is

$$\lambda^1(t) = -.25718\theta_0 t^3 + .77155\theta_0 t^2 - 1.0t - .514\,37\theta_0 + 1.0 \tag{4.208}$$

from which we find

$$\begin{aligned} \nabla_u J\left(u^1\right) = &-.25718\theta_0 t^3 + .77155\theta_0 t^2 - 1.0t - .51437\theta_0 + 1.0 \\ &+ 1.5431\theta_0 t - 1.5431\theta_0 \end{aligned}$$

Consequently

$$\begin{aligned} u^2(t) &= u^1(t) - \theta_1 \left[\nabla_u J\left(u^0\right)\right] \\ &= (1-\theta_1)(1.5431\theta_0 t - 1.5431\theta_0) \\ &\quad -\theta_1 \left[-.25718\theta_0 t^3 + .77155\theta_0 t^2 - 1.0t - .51437\theta_0 + 1.0\right] \end{aligned}$$

Note that although we started with a point estimate for the optimal control we now have a nonlinear function of time as our approximate optimal control. Note also that we have performed no explicit time discretizations in arriving at this approximation; that is, we have carried out the iterations of the algorithm in the appropriate function space. It is also important to recognize that the two-pointedness of the boundaries has also been completely overcome because the *algorithm always permits the adjoint and state dynamics to be solved separately*.

Let us pick

$$\begin{aligned} \theta_0 &= .76158 \\ \theta_1 &= .01 \end{aligned}$$

so that

$$\begin{aligned} u^2(t) = {} & 1.9586 \times 10^{-3} t^3 - 5.876 \times 10^{-3} t^2 \\ & + 1.1734t - 1.1695 \end{aligned} \tag{4.209}$$

As a check, let us compare expression (4.209), which is the solution obtained by steepest descent (SD) for $k = 2$, to the two-point boundary-value problem (TPBVP) solution obtained in Chapter 3 for the identical problem. That comparison is contained in the following table:

t	$u(t) : TPBVP$	$u(t) : SD, k = 2$
0.0	-1.1752	-1.1695
.25	$-.8223$	$-.8765$
.5	$-.5211$	$-.5840$
.75	$-.2526$	$-.2919$
1.0	$2.46 \times 10^{-5} \approx 0$	$-1.74 \times 10^{-5} \approx 0$

4.7.5 Example of the Gradient Projection Algorithm

Consider the following optimal control problem similar to the problem employed in the example of Section 4.7.4, recalling $u \in L^2[a,b]$ and $x \in \mathcal{H}^1[a,b]$:

$$\min J = \int_0^1 \frac{1}{2}\left(x^2 + u^2\right) dt \tag{4.210}$$

subject to

$$\frac{dx}{dt} = u \quad (\lambda) \tag{4.211}$$

$$x(0) = 2.5 \tag{4.212}$$

$$-1 \leq u \leq 1 \tag{4.213}$$

Let us apply the projected gradient algorithm to this problem. We know from Section 4.7.4 that

$$H = \frac{1}{2}\left(x^2 + u^2\right) + \lambda u \tag{4.214}$$

$$-\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = x \tag{4.215}$$

$$\lambda(1) = 0 \tag{4.216}$$

We also know that

$$\nabla_u J(u) = \frac{\partial H}{\partial u} = \lambda + u$$

We recall that the minimum norm projection onto $\Omega = \{u : a \leq u \leq b\}$ is

$$P_\Omega(v) = \arg\{\min \|v - y\| : y \in \Omega\} = [v]_a^b \tag{4.217}$$

where

$$[v]_a^b = \begin{cases} b & \text{if} \quad v > b \\ v & \text{if } a \leq u \leq b \\ a & \text{if} \quad v < a \end{cases} \tag{4.218}$$

So the iterative procedure that is the steepest descent algorithm in continuous time has the following structure:

1. guess $u^0(t)$
2. calculate $x^0(t)$ using $u^0(t)$ in (4.211),
3. calculate $\lambda^0(t)$ using $x^0(t)$ in (4.215)
4. calculate

$$d^0 = -\nabla J\left(u^0\right) = -\lambda^0 - u^0$$

5. iterate according to

$$u^1(t) = P_\Omega\left[u^0(t) + \theta_0 d^0(t)\right]$$

$$= \left[u^0(t) + \theta_0 d^0(t)\right]_{-1}^{+1}$$

6. repeat for subsequent iterations

The following calculations are grouped by iteration number k:

$\boxed{k=0}$: Pick the following initial feasible solution

$$u^0 = 0.5e^t - 1.5 \tag{4.219}$$

Our indicated choice of $u^0(t)$, namely expression (4.219), has the consequence that the state dynamics are

$$\frac{dx}{dt} = 0.5e^t - 1.5$$
$$x(0) = 2.5$$

The solution of the state dynamics for the current iteration is

$$x^0(t) = 0.50e^t - 1.50t + 2$$

Consequently the adjoint dynamics are

$$\frac{d\lambda}{dt} = -x^0(t) = -0.50e^t + 1.50t - 2$$
$$\lambda(1) = 0$$

The solution of the adjoint dynamics for the current iteration is

$$\lambda^0 = 0.75t^2 - 0.5e^t - 2.0t + 2.61$$

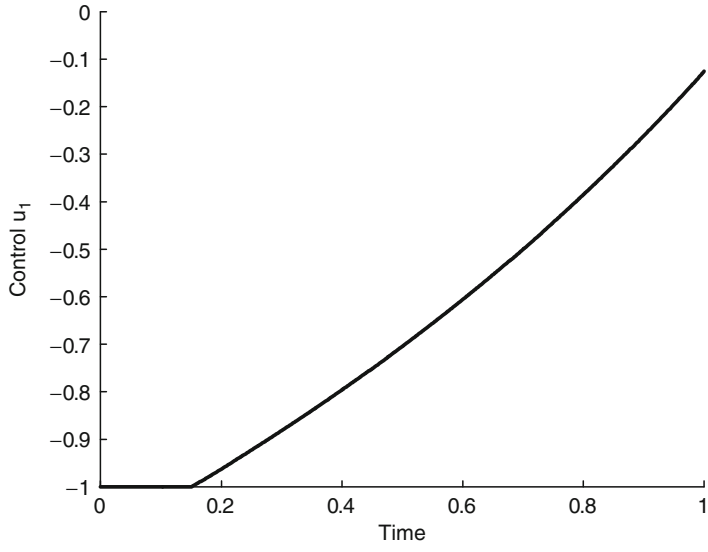
This leads directly to the descent direction

$$\begin{aligned} d^0 &= -\nabla_u J\left(u^0\right) = -\left(\lambda^0 + u^0\right) \\ &= -\left(0.75t^2 - 2.0t + 1.11\right) \\ &= -0.75t^2 + 2.0t - 1.11 \end{aligned}$$

Setting $\theta_0 = 0.1$, we have

$$\begin{aligned} u^1 &= \left[u^0 + \theta_0 d^0\right]_{-1}^{+1} \\ &= \left[0.2t + 0.5e^t - 0.075\,t^2 - 1.61\right]_{-1}^{+1} \\ &= \min\left(\max\left(-1, 0.2t + 0.5e^t - 0.075\,t^2 - 1.61\right), 1\right) \end{aligned}$$

whose plot is:



We see that there is a critical instant of time $t_b = 0.152$ after which the constraint $u^1 \geq -1$ ceases to bind; that time is a solution of the following equation:

$$0.2t + 0.5e^t - 0.075\,t^2 - 1.61 = -1$$

Thus, we have

$$u^1 = \begin{cases} -1 & \text{for} \quad t \in [0.0, 0.152] \\ 0.2t + 0.5e^t - 0.075\,t^2 - 1.61 & \text{for} \quad t \in (0.152, 1.0] \end{cases}$$

$\boxed{k = 1}$: For $t \in [0.0, 0.152]$, we have

$$\begin{aligned} u^1 &= -1 \\ x^1 &= -t + 1.54 \\ \lambda^1 &= 0.50t^2 - 1.54t + 1.04 \end{aligned}$$

Setting $\theta_1 = 0.1$, we have

$$\begin{aligned} d^1 &= -\nabla_u J\left(u^1\right) = -\left(\lambda^1 + u^1\right) \\ &= -\left(0.50t^2 - 1.54t + 0.04\right) \\ &= -0.50t^2 + 1.54t - 0.04 \end{aligned}$$

We update the current control to obtain

$$\begin{aligned} u^2 &= \left[u^1 + \theta_1 d^1\right]_{-1}^{+1} \\ &= \left[-0.05t^2 + 0.15t - 1.00\right]_{-1}^{+1} \\ &= \min\left(\max\left(-1, -0.05t^2 + 0.15t - 1.00\right), 1\right) \end{aligned}$$

For $t \in [0.0, 0.152]$, we find $u^2 = 0$. However, for $t \in (0.152, 1.0]$ we find that

$$\begin{aligned} \frac{dx}{dt} &= u^1 = 0.2t + 0.5e^t - 0.075\,t^2 - 1.61 \\ x(0) &= 2.5 \end{aligned}$$

Solving the above initial-value problem, we obtain

$$x^1(t) = 0.5e^t - 1.61t + 0.1t^2 - 0.025\,t^3 + 2.0$$

with the corresponding adjoint dynamics

$$\begin{aligned} \frac{d\lambda}{dt} &= -0.5e^t + 1.61t - 0.1t^2 + 0.025\,t^3 - 2.0 \\ \lambda(1) &= 0 \end{aligned}$$

Consequently,

$$\lambda^1(t) = 0.81t^2 - 0.5e^t - 2.0t - 0.033t^3 + 0.006t^4 + 2.58$$

The descent direction is given by

$$\begin{aligned} d^1 &= -\nabla_u J\left(u^1\right) = -\left(\lambda^1 + u^1\right) \\ &= -0.006t^4 + 0.033t^3 - 0.735t^2 + 1.8t - 0.97 \end{aligned}$$

Setting $\theta_1 = 0.1$, we obtain the following expression for u^2 when $t \in (0.1522, 1.0]$:

$$u^2 = \left[u^1 + \theta_1 d^1\right]_{-1}^{+1} = \left[0.38t + 0.5e^t - 0.15t^2 - 1.71\right]_{-1}^{+1}$$

For the argument of the above function a critical point for which $u^2 = -1$ is

$$t_b \approx 0.229$$

Therefore, our second iterate u^2 for all $t \in [0, 1]$ is

$$u^2 = \begin{cases} -1 & \text{for} \quad t \in [0.0, 0.229] \\ 0.38t + 0.5e^t - 0.15t^2 - 1.71 & \text{for} \quad t \in (0.229, 1.0] \end{cases}$$

Continuing in a similar manner one obtains after ten iterations the result

$$u^{10} = \begin{cases} -1 & \text{for} \quad t \in [0.0, 0.389] \\ 1.3027t + 0.50e^t - 0.63t^2 - 2.15 & \text{for} \quad t \in (0.389, 1.0] \end{cases}$$

If the algorithm is implemented on a computer, so that higher precision may be easily employed, the following table of iterations is obtained:

k	$u^k(t)$ for $t \in (t_k, 1)$	$(t_k, 1)$	Δ_k
0	$0.5e^t - 1.5$	–	–
1	$0.2t + 0.5e^t - 0.075\,t^2 - 1.6109$	$(0.1522, 1.0]$	0.11
2	$0.38t + 0.5e^t - 0.14805t^2 - 1.7079 + \ldots$	$(0.2289, 1.0]$	0.09
3	$0.542t + 0.5e^t - 0.21864t^2 - 1.7927 + \ldots$	$(0.2766, 1.0]$	0.05
4	$0.6878t + 0.5e^t - 0.28641t^2 - 1.8670 + \ldots$	$(0.3093, 1.0]$	0.035
5	$0.81902t + 0.5e^t - 0.35112t^2 - 1.932 + \ldots$	$(0.3328, 1.0]$	0.028
6	$0.93712t + 0.5e^t - 0.41261t^2 - 1.9890 + \ldots$	$(0.3504, 1.0]$	0.022
7	$1.0434t + 0.5e^t - 0.47080t^2 - 2.0389 + \ldots$	$(0.3638, 1.0]$	0.017
8	$1.1391t + 0.5e^t - 0.52567t^2 - 2.0827 + \ldots$	$(0.3743, 1.0]$	0.014
9	$1.2252t + 0.5e^t - 0.57724t^2 - 2.1211 + \ldots$	$(0.3826, 1.0]$	0.012
10	$1.3027t + 0.5e^t - 0.62558t^2 - 2.1548 + \ldots$	$(0.3892, 1.0]$	0.009

where $\Delta_k = \left|u^{k+1} - u^k\right|$.

4.7.6 Penalty Function Example

Consider the following problem where $u \in L^2[a, b]$ and $x \in \mathcal{H}^1[a, b]$:

$$\min J = \int_0^1 \frac{1}{2} x^2 dt \tag{4.220}$$

subject to

$$\frac{dx}{dt} = u \quad (\lambda) \tag{4.221}$$

$$x(0) = 1 \tag{4.222}$$

$$u \geq -1 \tag{4.223}$$

We need to form a penalty function for the single constraint

$$g(u) = -u - 1 \leq 0$$

Therefore, we seek to solve

$$\min J = \int_0^1 \left\{ \frac{1}{2}\left(x^2 + u^2\right) + \frac{1}{2}\left[\rho \max(0, g)\right]^2 \right\} dt \tag{4.224}$$

$$= \int_0^1 \left\{ \frac{1}{2}\left(x^2 + u^2\right) + \frac{1}{2}\left[\rho \max(0, -u - 1)\right]^2 \right\} dt \tag{4.225}$$

subject to

$$\frac{dx}{dt} = u \quad (\lambda) \tag{4.226}$$

$$x\,(0) = 1 \tag{4.227}$$

as $\rho \longrightarrow +\infty$. The augmented Hamiltonian is

$$H = \frac{1}{2}\left(x^2\right) + \frac{1}{2}\left[\rho \max(0, -u - 1)\right]^2 + \lambda u$$

The optimality conditions are

$$-\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} = x \tag{4.228}$$

$$\lambda\,(1) = 0 \tag{4.229}$$

$$u = \arg\left[\frac{\partial}{\partial u} H = 0\right] \tag{4.230}$$

$$\Longrightarrow \frac{\partial}{\partial u}\left[\frac{1}{2}\left(x^2\right) + \frac{1}{2}\left[\rho \max(0, -u - 1)\right]^2 + \lambda u\right] = 0 \tag{4.231}$$

Note that (4.231) reduces to

$$\rho\left[\max(0, -u - 1)\right](-1) + \lambda = 0$$

Hence, either

$$\lambda = 0$$

or

$$\rho\,(u + 1) + \lambda = 0$$

From the last equation, we obtain

$$u^* = \lim_{\rho \to \infty}\left[-1 - \frac{\lambda}{\rho}\right] = -1$$

The optimal state obeys

$$x^*(t) = 1 - t$$

The optimal adjoint variable is

$$\lambda^*(t) = \frac{1}{2}t^2 - t + \frac{1}{2}$$

4.8 Exercises

1. If $V = \Re^n$, what is its dual V^*?
2. If $V = L^2[a, b]$, what is its dual V^*?
3. Prove that any Hilbert space is reflexive.
4. Compare and contrast the notion of a variation with the notion of a G-derivative of a functional.
5. In the proof of Theorem 4.19, we make use of condition (4.112), repeated here for convenience

   $$\eta(\varepsilon) \longrightarrow 0_+ \Longleftrightarrow \varepsilon \longrightarrow 0_+$$

   Show that this property is assured by Lipschitz continuity as the notion is stated in Definition 4.19.
6. Extend the derivation in Section 4.3 of necessary conditions for continuous-time optimal control problems to include terminal costs and pure control constraints.
7. Solve this optimal control problem:

   $$\min J(u) = \int_0^1 \frac{1}{2}u^2 dt$$

   subject to

   $$\begin{aligned} \frac{dx}{dt} &= x + u(t - \tau) \\ 0 &\leq u \leq 1 \\ x(0) &= 1 \end{aligned}$$

   where $\tau = a + bx$ is a time shift.
8. Prove or disprove: linear functionals in a Hilbert space satisfy the Kuhn-Tucker constraint qualification in Definition 4.39.
9. Using the proof of Theorem 4.18 as a guide, extend the necessary conditions to handle equality constraints and give a proof that your conditions are in fact necessary.
10. Prove Lemma 4.3 related to coerciveness.
11. State as a theorem conditions that make (4.52) a sufficient as well as a necessary condition; formally prove your theorem.

12. Give an expanded statement of the penalty function algorithm (4.189), wherein individual steps are enumerated and a stopping test is expressed.
13. Discuss the challenges surrounding extension of the methods of this chapter to consider variable step sizes for mathematical programs in infinite dimensional spaces. Include a brief description of how these challenges might be overcome.
14. Solve the following optimal control problem using the steepest descent algorithm:

$$\min J(u) = \int_0^{10} \frac{1}{2}\left(x^2 - u\right) dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= x - u \\ x\,(0) &= 1 \end{aligned}$$

15. Solve the following optimal control problem using the gradient projection algorithm:

$$\min J(u) = \int_0^5 \frac{1}{2} x^2 dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= x - u \\ -1 \leq u &\leq 1 \\ x\,(0) &= 1 \end{aligned}$$

16. Solve the following optimal control problem using a penalty function:

$$\min J(u) = \int_0^5 \frac{1}{2} x^2 dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= u_2 - u_1 \\ (u_1)^2 + (u_2)^2 &\leq 2 \\ x\,(0) &= 1 \end{aligned}$$

List of References Cited and Additional Reading

Apostol, T. M. (1974). *Mathematical Analysis*. Reading, MA: Addison-Wesley.
Arrow, K. J. and A. C. Enthoven (1961). Quasi-concave programming. *Econometrica 29*, 779–800.
Bazaraa, M., H. Sherali, and C. Shetty (1993). *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley.

Berge, C. (1963). *Topological Spaces*. London: Oliver and Boyd.
Berman, A. (1973). *Cones, Matrices, and Mathematical Programming*. New York: Springer-Verlag.
Bressan, A. and B. Piccoli (2007). *An Introduction to the Mathematical Theory of Control*. Springfield, MO: American Institute of Mathematical Sciences.
Bryson, A. E. and Y. C. Ho (1975). *Applied Optimal Control*. New York: Hemisphere.
Budelis, J. J. and A. E. Bryson (1970). Some optimal control results for differential-difference systems. *IEEE Transactions on Automatic Control 15*, 237–241.
Canon, M., C. Cullum, and E. Polak (1970). *Theory of Optimal Control and Mathematical Programming*. New York: McGraw-Hill.
Debnath, L. and P. Mikusinski (1999). *Introduction to Hilbert Spaces with Applications*. London: Academic Press.
Fenchel, W. (1953). *Convex Cones, Sets, and Functions*. Princeton: Princeton Univesity Press.
Girsanov, I. V. (1972). *Lectures on Mathematical Theory of Extremum Problems*. New York: Springer-Verlag.
Guignard, M. (1969). Generalized kuhn-tucker conditions for mathematical programming problems in a banach space. *SIAM Journal of Control 7*, 232–241.
Lay, S. R. (1982). *Convex Sets and Their Applications*. New York: John Wiley.
Luenberger, D. G. (1966). A generalized maximum principle. In A. Lavi and T. P. Vogl (Eds.), *Recent Advances in Optimization Techniques*. New York: John Wiley.
Luenberger, D. G. (1969). *Optimization by Vector Space Methods*. New York: John Wiley.
Minoux, M. (1986). *Mathematical Programming: Theory and Algorithms*. New York: John Wiley.
Pervozvanskiy, A. A. (1967). Relationship between the basic theorems of mathematical programming and the maximum principle. *Engineering Cybernetics 6*, 11.
Polak, E. (1973, April). An histrical survey of computational methods in optimal control. *SIAM Review 15*(2), 553–584.
Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko (1962). *The Mathematical Theory of Optimal Processes*. New York: Interscience.
Ritter, K. (1967). Duality for nonlinear programming in a banach space. *SIAM Journal on Applied Mathematics 15*(2), 294–302.
Rockafellar, R. T. (1970). *Convex Analysis*. Princeton: Princeton University Press.
Rudin, W. (1987). *Real and Complex Analysis*. New York: McGraw-Hill.
Varaiya, P. (1967). Nonlinear programming in banach space. *SIAM Journal on Applied Mathematics 15*(2), 284–293.
Varaiya, P. P. (1965). Nonlinear programming and optimal control. Technical report, ERL Tech. Memo M-129. University of California, Berkeley.
Wouk, A. (1979). *A Course of Applied Functional Analysis*. New York: John Wiley.

Chapter 5
Finite Dimensional Variational Inequalities and Nash Equilibria

In this chapter, we lay the foundation for turning our focus from dynamic optimization, which has been the subject of preceding chapters, to the notion of a dynamic game. To fully appreciate the material presented in subsequent chapters, we must in the present chapter review some of the essential features of the theory of finite-dimensional variational inequalities and static noncooperative mathematical games. Today many economists and engineers are exposed to the notion of a game-theoretic equilibrium that we study in this chapter, namely *Nash equilibrium*. Yet, the relationship of such equilibria to certain nonextremal problems known as fixed-point problems, variational inequalities and nonlinear complementarity problems is not widely understood. It is the fact that, as we shall see, Nash and Nash-like equilibria are related to and frequently equivalent to nonextremal problems that makes the computation and qualitative investigation of such equilibria so tractable. Although the static games discussed in this chapter are really steady states of dynamic games, we are, for the most part, indifferent in this chapter to any underlying dynamics. We also comment that readers familiar with finite-dimensional variational inequalities and static Nash games may wish to skip this chapter.

The following is a preview of the principal topics covered in this chapter:

Section 5.1: Some Basic Notions. We begin this chapter by introducing the distinction between games in normal form and games in extensive form.

Section 5.2: Nash Equilibria and Normal Form Games. In this section, we define a Nash equilibrium and present the corresponding noncooperative game in normal form for which it is a solution.

Section 5.3: Some Related Nonextremal Problems. In this section, we introduce finite-dimensional, fixed point, complementarity, and variational inquality problems, as well as relationships among them.

Section 5.4: Sensitivity Analysis of Variational Inequalities. Because finite-dimensional variational inequalities are encountered as approximations of certain infinite-dimensional noncooperative games in subsequent chapters, we describe how to conduct sensitivity analysis of them.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_5, © Springer Science+Business Media, LLC 2010

Section 5.5: The Diagonalization Algorithm. In this section, we present the widely used diagonalization method for solution of finite-dimensional variational inequalities. We caution that it may fail to converge.

Section 5.6: Gap Function Methods for $VI(F, \Lambda)$. In this section, we present a number of so-called gap functions for finite-dimensional variational inequalities. The minimization of a gap function is observed to yield a solution of the underlying variational inequality.

Section 5.7: Other Algorithms for $VI(F, \Lambda)$. In this section, we present a brief survey of other solution methods for variational inequalities and noncooperative Nash games.

Section 5.8: Computing Network User Equilibria. The concluding section of this chapter illustrates how the well-known user equilibrium problem for road networks may be solved by a fixed-point algorithm.

5.1 Some Basic Notions

A mathematical game is a mathematical representation of some form of competition among agents or "players" of the game. Most mathematical games have rules of play, agent-specific utilities or payoffs, and a notion of solution. These may be expressed in two fundamental ways: the so-called extensive form and the normal form. A *game in extensive form* is a presentation, usually via a table or a decision tree, of all possible sequences of decisions that can be made by the game's players. This presentation is, by its very nature, exhaustive and potentially tedious or even impossible for large games involving multiple players and numerous decisions. By contrast a *game in normal form* is expressed via mappings, equations, inequalities, and extremal principles. As such, large normal form games are potentially much more computationally tractable, since they may draw upon the computational methods of mathematical programming and optimal control theory, as well as general variational methods.

5.2 Nash Equilibria and Normal Form Games

The best understood and most widely used mathematical games are noncooperative games, wherein game players, also called agents, act selfishly. A noncooperative mathematical game in normal form uses equations, inequalities, and extremal principles to describe competition among agents – who are intrinsically in conflict and do not collude – informed by some notion of utility and acting according to rules known by the agents of the game. We are especially interested in a notion of solution of noncooperative games known as a *Nash equilibrium* (named after John Forbes Nash, who proposed it). A set of actions undertaken by the noncooperative agents of interest is a Nash equilibrium if each agent knows the equilibrium strategies of

the other agents, and no agent has anything to gain by unilaterally changing his/her own strategy. In particular, if no agent can benefit by changing his/her strategy while the other agents keep theirs unchanged, then the current set of strategy choices and the corresponding payoffs constitute a Nash equilibrium. As such, finding the Nash equilibrium of a noncooperative game in normal form is not generally equivalent to a single optimization problem, but is, rather, naturally articulated as a family of coupled optimization problems. We will learn how, for certain assumptions, those coupled optimization problems may be expressed as so-called *nonextremal problems*. Certain nonextremal problems have a structure that makes them quite amenable to analysis and solution. For our purposes in this chapter, the nonextremal problems known as fixed-point problems, variational inequality problems, and nonlinear complementarity problems are the most important; below, we define each in turn.

The following definition will apply:

Definition 5.1. *Nash equilibrium. Suppose there are N agents, each of which chooses a feasible strategy vector x^i from the strategy set Ω_i which is independent of the other players' strategies. Furthermore, every agent $i \in [1, N] \subseteq \mathcal{I}_{++}$ has a cost (disutility) function $\Theta_i(x) : \Omega \longrightarrow \Re^1$ that depends on all agents' strategies where*

$$\Omega = \prod_{i=1}^{N} \Omega_i$$
$$x = \left(x^i : i = 1, \ldots, N\right)$$

Every agent $i \in [1, N]$ seeks to solve the problem

$$\min \Theta_i(x^i, x^{-i}) \quad \text{s.t.} \quad x^i \in \Omega_i \tag{5.1}$$

for each fixed yet arbitrary non-own tuple

$$x^{-i} = \left(x^j : j \neq i\right)$$

A Nash equilibrium is a tuple of strategies x, one for each agent, such that each x^i solves the mathematical program (5.1), and is denoted as $NE(\Theta, \Omega)$.

In other words no agent may lower his/her cost (disutility) by unilaterally altering his/her strategy.

When the strategy set of any agent $i \in [1, N]$ depends on non-own strategies x^j where $j \neq i$, extension of the definition of a Nash equilibrium is called a generalized Nash equilibrium. That is, we have the following definition:

Definition 5.2. *Generalized Nash equilibrium. Suppose there are N agents, each of which chooses a feasible strategy vector x^i from the strategy set $\Omega_i(x)$ that depends on the strategies of all agents where*

$$x = \left(x^i : i = 1, \ldots, N\right)$$

Furthermore, every agent $i \in [1, N] \subseteq \mathcal{I}_{++}$ *has a cost (disutility) function* $\Theta_i(x) : \Omega(x) \longrightarrow \Re^1$ *that depends on all agents' strategies where*

$$\Omega(x) = \prod_{i=1}^{N} \Omega_i(x)$$

Every agent $i \in [1, N]$ *seeks to solve the problem*

$$\min \Theta_i(x^i, x^{-i}) \quad \text{s.t.} \quad x^i \in \Omega_i(x) \tag{5.2}$$

for each fixed yet arbitrary non-own tuple

$$x^{-i} = \left(x^j : j \neq i\right)$$

A generalized Nash equilibrium is a tuple of strategies x*, one for each agent, such that each* x^i *solves the mathematical program (5.1), and is denoted as* $GNE(\Theta, \Omega)$.

5.3 Some Related Nonextremal Problems

We now define the fixed-point problem:

Definition 5.3. *Fixed-point problem. Given a nonempty set* $\Lambda \subseteq \Re^n$ *and a function* $F : \Lambda \to \Lambda$*, the fixed-point problem* $FPP(F, \Lambda)$ *is to find a vector* y *such that*

$$\left.\begin{array}{c} y \in \Lambda \\ y = F(y) \end{array}\right\} \quad FPP(F, \Lambda) \tag{5.3}$$

If $\Lambda \subseteq \Re^1$, then $FPP(G, \Lambda)$ seeks to find a point where the graph of F crosses the 45-degree line.

It will also be useful to define an extension of $FPP(F, \Lambda)$ which employs the notion of a minimum norm projection. The minimum norm projection of the vector $v \in \Re^n$ onto the set Λ is denoted as $P_\Lambda[v]$ and has the following definition:

Definition 5.4. *Minimum norm projection.* $P_\Lambda[v]$*, the minimum norm projection of the vector* $v \in \Re^n$ *onto the set* $\Lambda \subseteq \Re^n$*, is the vector*

$$y = \arg\left\{\min_x \|v - x\| : x \in \Lambda\right\}$$

The fixed-point problem with projection is:

Definition 5.5. *Fixed-point problem based on projection. Given a nonempty set* $\Lambda \subseteq \Re^n$*, a fixed vector* $v \in \Lambda$*, and a function* $F : \Lambda \to \Lambda$*, the fixed-point problem based on the minimum norm projection* $FPP_{\min}(F, \Lambda)$ *is to find a vector* y *such that*

$$\left.\begin{array}{c} y \in \Lambda \\ y = P_{\Lambda}\left[y - F(y)\right] \end{array}\right\} \quad FPP_{\min}(F, \Lambda) \tag{5.4}$$

That is, the solution of $FPP_{\min}(F, \Lambda)$ is

$$y = \arg\left\{\min_{x} \|v - F(y) - x\| : x \in \Lambda\right\} \tag{5.5}$$

by virtue of the definition of the minimum norm operator.

Next, we define the variational inequality problem:

Definition 5.6. *Variational inequality. Given a nonempty set $\Lambda \subseteq \Re^n$ and a function $F : \Lambda \to \Re^n$, the variational inequality problem $VI(F, \Lambda)$ is to find a vector y such that*

$$\left.\begin{array}{c} y \in \Lambda \\ [F(y)]^T (x - y) \geq 0 \quad \forall\ x \in \Lambda \end{array}\right\} \quad VI(F, \Lambda) \tag{5.6}$$

Geometrically, a vector y is a solution of $VI(F, \Lambda)$ if and only if $F(y)$ forms an acute or right angle with all feasible vectors emanating from y.

Finally, we define the nonlinear complementarity problem:

Definition 5.7. *Nonlinear complementarity problem. Given a (nonlinear) function $F : \Re^n \to \Re^n$, the nonlinear complementarity problem $NCP(F)$ is to find a vector y such that*

$$\left.\begin{array}{c} [F(y)]^T \cdot y = 0 \\ F(y) \geq 0 \\ y \geq 0 \end{array}\right\} \quad NCP(F) \tag{5.7}$$

Geometrically, a vector y is a solution of $NCP(F)$ if and only if y is nonnegative, $F(y)$ is nonnegative, and $F(y)$ is orthogonal to y. Alternatively, a vector y is a solution of $NCP(F)$ if and only if all elements of y are nonnegative, all elements of $F(y)$ are nonnegative, and for each positive element of y, denoted by y_i, $F_i(y)$ is zero (and vice versa).

5.3.1 Nonextremal Problems and Programs

Let us begin our analysis of nonextremal problems by stating and proving a key result that relates variational inequalities to mathematical programs. That result is

Theorem 5.1. *Variational inequality global optimality condition. A necessary and sufficient condition for $y \in \Lambda \subseteq \Re^n$ to be a global optimum of the mathematical program*

$$\left.\begin{array}{l} \min\ Z(x) \\ \text{s.t. } x \in \Lambda \end{array}\right\} \quad NLP(F, \Lambda) \tag{5.8}$$

when Λ is non-empty and convex and $Z(x)$ is convex and differentiable for all $x \in \Lambda$, is the variational inequality

$$[\nabla Z(x)]^T (x - y) \geq 0 \quad \forall x \in \Lambda \tag{5.9}$$

Proof. The proof is in two parts:

(i) [(5.9) $\Longrightarrow$ (5.8)] The well-known property that any tangent to a differentiable convex function underestimates that function is expressed for the present case as

$$Z(x) \geq Z(y) + [\nabla Z(y)]^T (x - y^*) \quad \forall x \in \Lambda \tag{5.10}$$

It is immediate from (5.10) that

$$Z(x) - Z(y^*) \geq [\nabla Z(y^*)]^T (x - y^*) \geq 0 \quad \forall x \in \Lambda \tag{5.11}$$

in light of the given (5.9), thereby establishing sufficiency.

(ii) [(5.8) $\Longrightarrow$ (5.9)] Necessity is established by observing that following any direction vector $(x-y)$ rooted at the global optimum y must lead to another feasible solution that increases the objective function of (5.8). That is, every $(x-y)$ must have a component in the direction of $\nabla Z(y)$, a circumstance ensured by (5.9). ■

5.3.2 *Kuhn-Tucker Conditions for Variational Inequalities*

The so-called Kuhn-Tucker necessary conditions for finite-dimensional mathematical programs, as we presented them in Chapter 2, express the gradient as a linear combination of binding constraints at optimality. Kuhn-Tucker type necessary conditions may also be developed for variational inequalities and Nash equilibria; these conditions are needed for a variety of applications in subsequent chapters as well as for the sensitivity analysis of variational inequalities discussed in the next section. Our development of Kuhn-Tucker conditions for variational inequalities parallels that in Tobin (1986) and depends on observing that $VI(F, \Lambda)$ requires

$$[F(x^*)]^T x \geq [F(x^*)]^T x^* \quad \forall x \in \Lambda \tag{5.12}$$

This last inequality is recognized as the definition of a constrained global minimum for the objective function $[F(x^*)]^T x$. That is, $VI(F, \Lambda)$ can be restated as the following mathematical program:

$$\min [F(x^*)]^T x \quad \text{s.t.} \quad x \in \Lambda \subseteq \Re^n \tag{5.13}$$

where

$$F(x) : \Re^n \longrightarrow \Re^n$$

Note carefully that (5.13) is of no real use for computation as it presumes knowledge of the solution $x^* \in \Lambda$.

In our development of Kuhn-Tucker conditions for variational inequalities, we will consider the feasible region Λ to be determined by equality and inequality constraints; that is

$$\Lambda = \{x \in \Re^n : h(x) = 0, g(x) \leq 0\} \tag{5.14}$$

where

$$h(x) : \Re^n \longrightarrow \Re^q$$
$$g(x) : \Re^n \longrightarrow \Re^m$$

We will further assume that the functions $F(x)$ and $g(x)$ are both continuous, $g(x)$ is differentiable on Λ, while $h(x)$ is linear affine on Λ. The key result is:

Theorem 5.2. *Necessary conditions for VI* (F, Λ). *Let*

$$x^* \in \Lambda = \{x \in \Re^n : h(x) = 0, g(x) \leq 0\}$$

be a solution of VI (F, Λ). *Further assume that* $F(x)$ *and* $g(x)$ *are continuous on* Λ, $g(x)$ *is differentiable on* Λ, *and* $h(x)$ *is linear affine on* Λ. *Then, if the gradients* $\nabla g_i(x^*)$ *for* i *such that* $g_i(x^*) = 0$ *together with the gradients* $\nabla h_i(x^*)$ *for* $i \in [1, q]$ *are linearly independent, there exist multipliers* $\pi \in \Re^m$ *and* $\mu \in \Re^q$ *such that*

$$F(x^*) + \left[\nabla g(x^*)\right]^T \pi + \left[\nabla h(x^*)\right]^T \mu = 0 \tag{5.15}$$
$$\pi^T g\left(x^*\right) = 0 \tag{5.16}$$
$$\pi \geq 0 \tag{5.17}$$

Proof. Observe that x^* also solves the nonlinear program

$$\min Z\left(x^*\right) \equiv \left[F\left(x^*\right)\right]^T x \quad \text{s.t.} \quad x \in \Lambda \subseteq \Re^n \tag{5.18}$$

The assumption of linear independence of the gradients of binding constraints is a sufficient condition for the Kuhn-Tucker constraint qualification to hold at x^*; therefore, the Kuhn-Tucker conditions for this mathematical program are

$$\nabla Z\left(x^*\right) + \pi^T \nabla g\left(x^*\right) + \mu^T \nabla h\left(x^*\right) = 0 \tag{5.19}$$
$$\pi^T g\left(x^*\right) = 0 \tag{5.20}$$
$$\pi \geq 0 \tag{5.21}$$

However

$$\nabla Z\left(x^*\right) = F\left(x^*\right) \tag{5.22}$$

so (5.19), (5.20), and (5.21) are equivalent to (5.15), (5.16), and (5.17). ■

Note that Theorem 5.2 can be strengthened by employing a weaker (less restrictive) constraint qualification.

We further comment that the variational inequality necessary conditions become sufficient if we stipulate that the inequality constraint functions $g_i(x)$ are convex. This observation is formalized in the next theorem:

Theorem 5.3. *Sufficient conditions for VI* (F, Λ). *Suppose the assumptions of Theorem 5.2 hold; the* $g_i(x)$ *for* $i \in [1, m]$ *are convex on* Λ*; and* $x^* \in \Lambda$, $\pi \in \Re^m$, $\mu \in \Re^q$ *satisfy (5.15), (5.16), and (5.17). Then* x^* *is a solution to VI* (F, Λ).

Proof. By the given of this theorem, the nonlinear program (5.18) is a convex mathematical program. Thus, (5.15), (5.16), and (5.17) are sufficient to conclude that x^* solves (5.18). Consequently

$$\left[F\left(x^*\right)\right]^T x \geq \left[F\left(x^*\right)\right]^T x^* \quad \forall x \in \Lambda \tag{5.23}$$

demonstrating that x^* solves $VI(F, \Lambda)$. ■

5.3.3 Variational Inequality and Complementarity Problem Generalizations

It is possible to generalize $VI(F, \Lambda)$ in a variety of ways. Two of those generalizations are defined below:

Definition 5.8. *Generalized variational inequality. Given a nonempty set* $\Lambda \subseteq \Re^n$ *and functions* $F : \Re^n \to \Re^n$ *and* $f : \Re^n \to \Re^n$, *the* generalized variational inequality problem $GVI(F, f, \Lambda)$ *is to find a vector* $y \in \Re^n$ *such that*

$$\left.\begin{array}{c} f(y) \in \Lambda \\ [F(y)]^T [f(x) - f(y)] \geq 0 \quad \forall f(x) \in \Lambda \end{array}\right\} \; GVI(F, \Lambda)$$

Definition 5.9. *Quasivariational inequality. Given a function* $F : \Re^n \to \Re^n$ *and the point-to-set mapping* $K(y)$ *such that* $K(y) \subset \Re^n$, *the* quasivariational inequality problem $QVI(\Lambda, F)$ *is to find a vector* $y \in K(y)$ *such that*

$$\left.\begin{array}{c} y \in K(y) \\ [F(y)]^T (x - y) \geq 0 \quad \forall x \in K(y) \end{array}\right\} \; QVI(F, \Lambda)$$

5.3.4 Relationships Among Nonextremal Problems

There is a variety of relationships among the various nonextremal problems we have defined. We formalize some of those relationships in the following results:

Lemma 5.1. *Nonlinear complementarity and variational inequality equivalence. If* $\Lambda = \Re^n_+$ *then the variational inequality problem* $VI(F, \Lambda)$ *is equivalent to the nonlinear complementarity problem* $NCP(F)$.

Proof. The proof is in two parts:

(i) [$NCP(F) \Longrightarrow VI(F, \Lambda)$] First we show that if y is a solution of $NCP(F)$ then it is also a solution to $VI(F, \Lambda)$. To do so, note that, if y is a solution to $NCP(F)$, then $F(y) \geq 0$. Therefore, $[F(y)]^T x \geq 0$ for all $x \in \Lambda$. Thus, since $[F(y)]^T y = 0$ for all $y \geq 0$, it follows that

$$[F(y)]^T x - [F(y)]^T y \geq 0 \tag{5.24}$$

$$\Longrightarrow \quad [F(y)]^T (x - y) \geq 0 \quad x, y \in \Lambda. \tag{5.25}$$

(ii) [$VI(F, \Lambda) \Longrightarrow NCP(F)$] Next we show that if y is a solution of $VI(F, \Lambda)$ then it is also a solution of $NCP(F)$. To do so, note that if y is a solution to $VI(F, \Lambda)$ then

$$[F(y)]^T (x - y) \geq 0 \quad \forall\, x \in \Lambda \tag{5.26}$$

$$\Longrightarrow [F(y)]^T (-y) \geq 0, \text{since } x = 0 \in \Lambda \tag{5.27}$$

$$\Longrightarrow [F(y)]^T y \leq 0 \tag{5.28}$$

Also note that $x = 2y \in \Lambda$ since $y \in \Lambda$. Thus,

$$[F(y)]^T (2y - y) \geq 0 \Longrightarrow [F(y)]^T y \geq 0. \tag{5.29}$$

However

$$[F(y)]^T y \geq 0 \text{ and } [F(y)]^T y \leq 0 \Longrightarrow [F(y)]^T y = 0 \tag{5.30}$$

By assumption

$$[F(y)]^T (x - y) = \sum_{i=1}^{n} F_i(y)(x_i - y_i) \geq 0 \quad \forall\, x \in \Lambda. \tag{5.31}$$

Now, suppose $F(y) \neq 0$ and $F(y) \not> 0$. Then, there exists a $j \in [1, n]$ such that $F_j(y) < 0$. So, pick $x_j = +\infty > 0$. Then, it is immediate that

$$\sum_i F_i(y)(x_i - y_i) < 0 \tag{5.32}$$

which is a contradiction of our supposition; hence $F(y) \geq 0$. Also, since $y \in \Lambda$, it is immediate that $y \geq 0$. Thus,

$$[F(y)]^T y = 0, \quad F(y) \geq 0, \qquad y \geq 0. \tag{5.33}$$

■

Lemma 5.2. *Fixed-point problem and variational inequality equivalence. The fixed-point problem based on the minimum norm projection* $FPP_{\min}(F, \Lambda)$ *is equivalent to the variational inequality problem VI* (F, Λ) *when* $\Lambda \subseteq \Re^n$ *is a convex set.*

Proof. By definition $FPP_{\min}(F, \Lambda)$ can be viewed as requiring the solution of

$$\min_x \|y - F(y) - x\| \quad \text{s.t.} \quad x \in \Lambda \tag{5.34}$$

The mathematical program (5.34) is equivalent to

$$\min_x \frac{1}{2}(y - F(y) - x)^T (y - F(y) - x) \equiv Z(x; y) \quad \text{s.t.} \quad x \in \Lambda \tag{5.35}$$

since the objective function of (5.35) is a monotonic transformation of the objective function of (5.34). By Theorem 5.1 a necessary and sufficient condition for $x^* \in \Lambda$ to be an optimal solution of (5.35) is

$$[\nabla_x Z(x; y)]^T \left(x - x^*\right) \geq 0 \quad \forall x \in \Lambda \tag{5.36}$$

or

$$(-1)(y - F(y) - x)^T \left(x - x^*\right) \geq 0 \quad \forall x \in \Lambda \tag{5.37}$$

Since by the given we are solving a fixed-point problem, we know that $y = x^*$; it is then immediate from (5.37) that

$$[F(y)]^T (x - y) \geq 0 \quad \forall x \in \Lambda$$

as required. ■

Observe that if the vector function $F(y)$ is replaced by $\eta F(y)$ where $\eta \in \Re^1_{++}$ the result is unchanged.

It is interesting to note that the following result also holds:

Theorem 5.4. *Nonlinear complementarity and variational inequality equivalence. There is a nonlinear complementarity problem that is equivalent to VI* (F, Λ) *provided VI* (F, Λ) *obeys a constraint qualification and*

$$\Lambda = \{x \geq 0 : g(x) \leq 0, h(x) = 0\} \subseteq \Re^n_+$$

is convex.

Proof. Taking

$$F(x) : \Re^n \longrightarrow \Re^n$$
$$g(x) : \Re^n \longrightarrow \Re^m$$
$$h(x) : \Re^n \longrightarrow \Re^q$$

define

$$\Psi(y) = \begin{pmatrix} F(x^*) + [\nabla g(x^*)]^T \lambda + [\nabla h(x^*)]^T \mu \\ -g(x^*) \\ h(x^*) \\ -h(x^*) \end{pmatrix} \in \Re^{n+m+2q} \tag{5.38}$$

and

$$y = \begin{pmatrix} x \\ \lambda \\ \gamma \\ \eta \end{pmatrix} \in \Re^{n+m+2q} \tag{5.39}$$

We know that $h(x^*) = 0$ is equivalent to

$$h(x^*) \leq 0 \text{ and } h(x^*) \geq 0 \tag{5.40}$$

The Kuhn-Tucker identity for $VI(F, \Lambda)$ is

$$F(x^*) + [\nabla g(x^*)]^T \lambda + [\nabla h(x^*)]^T \mu = \rho \tag{5.41}$$

while the pertinent complementary slackness conditions are

$$\lambda^T g(x^*) = 0 \tag{5.42}$$
$$-g(x^*) \geq 0 \tag{5.43}$$
$$\lambda \geq 0 \tag{5.44}$$

and

$$\rho^T x^* = 0 \tag{5.45}$$
$$x^* \geq 0 \tag{5.46}$$
$$\rho \geq 0 \tag{5.47}$$

and

$$\gamma^T h(x^*) = 0 \tag{5.48}$$
$$h(x^*) \geq 0 \tag{5.49}$$
$$\gamma \geq 0 \tag{5.50}$$

and

$$\eta^T [-h(x^*)] = 0 \tag{5.51}$$
$$-h(x^*) \geq 0 \tag{5.52}$$
$$\eta \geq 0 \tag{5.53}$$

It is immediate from complementary slackness and the Kuhn-Tucker identity that

$$\left(F\left(x^*\right)+\left[\nabla g\left(x^*\right)\right]^T\lambda+\left[\nabla h\left(x^*\right)\right]^T\mu\right)^T x^*=0 \tag{5.54}$$

$$F\left(x^*\right)+\left[\nabla g\left(x^*\right)\right]^T\lambda+\left[\nabla h\left(x^*\right)\right]^T\mu\geq 0 \tag{5.55}$$

$$x\geq 0 \tag{5.56}$$

The nonlinear complementarity problem

$$\begin{aligned}\left[\Psi\left(y\right)\right]^T y&=0\\ \Psi\left(y\right)&\geq 0\\ y&\geq 0\end{aligned}$$

follows, if we employ definitions (5.38) and (5.39). Because of convexity the Kuhn-Tucker conditions are also sufficient. Hence, the two problems are equivalent. ■

Note that variational inequalities are more "general" than nonlinear programs. To see this, consider the nonlinear program

$$\left.\begin{array}{ll}\min & \oint_0^x F(z)dz\\ \text{s.t.} & x\in\Lambda\end{array}\right\}\quad NLP\left[\oint F(z)dz,\Lambda\right] \tag{5.57}$$

where $\oint$ denotes a line integral which must be well defined and yield a single valued function on Λ for (5.57) to be meaningful. For this program, we have the following result:

Lemma 5.3. *Nonlinear program and variational inequality equivalence. Let* Λ *be convex and take*

$$\oint_0^x F(z)dz$$

to be single valued and strictly convex on Λ. *Then, the nonlinear program* $NLP\left[\oint F(z)dz,\Lambda\right]$ *is equivalent to the variational inequality* $VI\left(F,\Lambda\right)$. *That is, the variational inequality* $VI\left(F,\Lambda\right)$ *is a necessary and sufficient condition for optimality of the nonlinear program* $NLP\left[\oint F(z)dz,\Lambda\right]$.

Proof. By Theorem 5.1 we know that a necessary and sufficient condition for optimality of $NLP\left[\oint F(z)dz,\Lambda\right]$ is

$$\left[\nabla_x\oint_0^x F(z)dz\right]^T_{x=y}(x-y)=\left[F(y)\right]^T(x-y)\geq 0,\quad\forall\, x\in\Lambda \tag{5.58}$$

which is $VI\left(F,\Lambda\right)$. ■

5.3.5 Variational Inequality Representation of Nash Equilibrium

Although the kinds of nonextremal problems introduced above are very interesting in their own right, they are perhaps most useful in the formulation of both network and nonnetwork game-theoretic equilibrium models. Loosely speaking, a system is in equilibrium when fluctuations have ceased. Thus, if we think of the function $G(y) - y$ as embodying the signals that guide how a system evolves over time, an equilibrium exists when the fixed-point problem $G(y) = y$ obtains. It is therefore no surprise that most equilibrium models can be formulated as fixed-point problems. In light of the connection between fixed-point and variational inequality problems that we have established, we fully expect that a Nash equilibrium in the sense of Definition 6.3 will be equivalent to a variational inequality under appropriate regularity conditions. In fact, the following result may be stated and proven:

Theorem 5.5. *Nash equilibrium equivalent to a variational inequality. The Nash equilibrium* $NE(\Theta, \Omega)$ *of Definition 5.1 is equivalent to the variational inequality* $VI(\nabla\Theta, \Omega)$ *the following regularity conditions hold: (1) each* $\Theta_i\,(x) : \Omega_i \longrightarrow \Re^1$ *is convex and continuously differentiable in* x^i; *and (2) each* Ω_i *is a closed convex subset of* $\Re^{n_i}$.

Proof. Each agent $i \in [1, N]$ seeks to solve

$$\min \Theta_i(x^i, x^{-i}) \quad \text{s.t.} \quad x^i \in \Omega_i \tag{5.59}$$

Because of convexity and differentiability, the variational inequality principle provides a necessary and sufficient condition for $y^i \in \Omega_i$ to be an equilibrium, namely

$$\nabla_i \Theta_i(y^i, y^{-i}) \left(x^i - y^i\right) \geq 0 \quad \forall x^i \in \Omega_i \;\; i \in [1, N] \tag{5.60}$$

where ∇_i denotes the gradient operator relative to x^i for $i \in [1, N]$. Concatenating the expressions (5.60) gives

$$\left[\nabla\Theta(y)\right]^T (x - y) \geq 0 \quad \forall x \in \Omega \tag{5.61}$$

which is recognized as $VI(\nabla\Theta, \Omega)$. Now suppose we are given (5.61); we may, for any arbitrary $i \in [1, N]$, select the tuple x to have y^j as its j^{th} subvector for every $j \neq i$. As a consequence of such choices, (5.61) yields the expressions (5.60). Thereby, the desired equivalency has been demonstrated. ■

5.3.6 User Equilibrium

In vehicular traffic science much effort has been devoted to modeling and computing a Nash-like equilibrium known as *user equilibrium*. This type of equilibrium is

a steady state flow pattern that is sometimes also called *user-optimized flow*. Traffic is said to achieve a user equilibrium when no traveler can change his/her route without experiencing greater travel delay or increased generalized cost (that includes consideration of the value of time).

To construct a model of user equilibrium we begin with a general network $G(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is a set of nodes and $\mathcal{A}$ is a set of arcs. We use $(i, j) \in \mathcal{W}$ to denote an origin-destination (OD) pair for which the origin is node i and the destination is node j while the set of all OD pairs is $\mathcal{W}$. There is a fixed travel demand T_{ij}, expressed in flow units, for each OD pair $(i, j) \in \mathcal{W}$. Furthermore, the minimum travel cost for OD pair $(i, j) \in \mathcal{W}$ is u_{ij}. The set of paths from node i to node j is denoted by $\mathcal{P}_{ij}$, while the unit cost of travel over path $p \in \mathcal{P}_{ij}$ is denoted by c_p. In addition, we denote the flow on path p by h_p. Letting $\mathcal{P}$ denote the set of all paths in the network, we are able to say the vector $h = \left(h_p : p \in \mathcal{P}\right) \geq 0$ is a user equilibrium when it obeys the following:

$$h_p > 0, \ p \in \mathcal{P}_{ij} \Longrightarrow c_p = u_{ij} \tag{5.62}$$

where

$$u_{ij} = \min_{p \in \mathcal{P}_{ij}} c_p$$

and flow conservation constraints are also enforced. The condition

$$c_p > u_{ij}, \ p \in \mathcal{P}_{ij} \Longrightarrow h_p = 0 \tag{5.63}$$

is automatically enforced and need not be separately articulated, as may be easily established by assuming $h_p > 0$ when $c_p > u_{ij}$ for $p \in \mathcal{P}_{ij}$. Using (5.62) a contradiction immediately results, verifying (5.63).

Usually path costs are taken to be additive in unit arc costs; that is

$$c_p = \sum_{a \in \mathcal{A}} \delta_{ap} c_a(f) \quad \forall p \in \mathcal{P} \tag{5.64}$$

where c_a is a unit cost function that reflects congestion by depending on the vector of arc flows $f = (f_a : a \in \mathcal{A})$ while f_a is the flow on each arc $a \in \mathcal{A}$. We will use the vectors

$$c = (c_a : a \in \mathcal{A})$$
$$C = (c_p : p \in \mathcal{P})$$

to denote the vector of arc costs and the vector of path costs, respectively. Also

$$\delta_{ap} = \begin{cases} 1 & if \ \text{arc } a \text{ belongs to path } p \\ 0 & if \ \text{arc } a \text{ does not belong to path } p \end{cases} \tag{5.65}$$

for each arc $a \in \mathcal{A}$ and path $p \in \mathcal{P}$. Arc flows are related to path flows according to

$$f_a = \sum_{p \in \mathcal{P}} \delta_{ap} h_p \quad \forall a \in \mathcal{A} \tag{5.66}$$

The relationships

$$T_{ij} - \sum_{p \in \mathcal{P}_{ij}} h_p = 0 \qquad \forall (i,j) \in \mathcal{W} \tag{5.67}$$

are the conservation of flow constraints. Clearly, then, the set

$$\Upsilon = \left\{ h \geq 0 : T_{ij} - \sum_{p \in \mathcal{P}_{ij}} h_p = 0 \quad \forall i, j \right\}$$

is the set of feasible flows from which the user equilibrium must be selected.

The user equilibrium problem we have desbribed above may be restated in a number of ways. One version is the following:

Definition 5.10. *User equilibrium with fixed demand. A user equilibrium* $UE(C, \Upsilon)$ *with fixed demand* $T = \left(T_{ij} : (i,j) \in \mathcal{W}\right)$ *is a flow pattern* $h \equiv \left(h_p : p \in \mathcal{P}\right)$ *such that*

$$\left(c_p - u_{ij}\right) h_p = 0 \qquad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij} \tag{5.68}$$

$$c_p - u_{ij} \geq 0 \qquad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij} \tag{5.69}$$

$$\sum_{p \in \mathcal{P}_{ij}} h_p - T_{ij} = 0 \qquad \forall (i,j) \in \mathcal{W} \tag{5.70}$$

$$h_p \geq 0 \qquad \forall p \in \mathcal{P} \tag{5.71}$$

where

$$u_{ij} = \min_{p \in \mathcal{P}_{ij}} c_p \quad \forall (i,j) \in \mathcal{W}$$

The system (5.68), (5.69), (5.70), and (5.71) looks quite similar to a nonlinear complementaity problem but is not. However, under the assumption of cost positivity, it is equivalent to a nonlinear complementarity problem:

Theorem 5.6. *User equilibrium as a nonlinear complementarity problem. Assume each arc cost* $c_a(f)$ *is strictly positive for all feasible flow and all* $a \in \mathcal{A}$. *Any pair* (h, u), *where* $h = \left(h_p : p \in \mathcal{P}\right)$ *and* $u = \left(u_{ij} : (i,j) \in \mathcal{W}\right)$, *is a user equilibrium* $UE(C, \Upsilon)$ *if it satisfies the following nonlinear complementarity problem:*

$$\left(c_p - u_{ij}\right) h_p = 0 \qquad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij} \tag{5.72}$$

$$c_p - u_{ij} \geq 0 \qquad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij} \tag{5.73}$$

$$\left(\sum_{q \in \mathcal{P}_{ij}} h_q - T_{ij} \right) u_{ij} = 0 \qquad \forall (i,j) \in \mathcal{W} \tag{5.74}$$

$$\left(\sum_{p \in \mathcal{P}_{ij}} h_p - T_{ij} \right) \geq 0 \qquad \forall (i,j) \in \mathcal{W} \tag{5.75}$$

$$h_p \geq 0 \qquad \forall p \in \mathcal{P} \tag{5.76}$$

Proof. That any solution of system (5.72), (5.73), (5.74), (5.75), and (5.76) is a solution of system (5.68), (5.69), (5.70), and (5.71) is seen by the following argument: if the desired relationship does not hold, the following chain of implications obtains:

$$\sum_{q \in \mathcal{P}_{ij}} h_q - T_{ij} > 0 \Longrightarrow u_{ij} = 0 \text{ and } \exists\, h_q > 0 \Longrightarrow c_q = u_{ij} = 0 \tag{5.77}$$

Clearly (5.77) violates the assumption of cost positivity; thereby we have established that any solution of (5.72), (5.73), (5.74), (5.75), and (5.76) is a user equilibrium. ■

It is an easy matter to show that a flow pattern is a user equilibrium if and only if it satisfies an appropriate variational inequality:

Theorem 5.7. *User equilibrium as a variational inequality. The flow pattern* $h^* = \left(h_p^* : p \in \mathcal{P}\right)$ *is a user equilibrium if and only if*

$$\left.\begin{array}{c} h^* \in \Upsilon \\ \sum_{p \in \mathcal{P}} c_p\left(h^*\right)\left(h_p - h_p^*\right) \geq 0 \qquad \forall h \in \Omega \end{array}\right\} \quad VI(c, \Omega) \tag{5.78}$$

where

$$\Upsilon = \left\{ h \geq 0 : T_{ij} - \sum_{p \in \mathcal{P}_{ij}} h_p = 0 \quad \forall (i,j) \in \mathcal{W} \right\} \tag{5.79}$$

Proof. The proof is in two parts:
(i) $[UE\,(C, \Upsilon) \Longrightarrow VI\,(C, \Upsilon)]$ Note that

$$c_p\left(h^*\right) \geq u_{ij}$$

for any $p \in \mathcal{P}_{ij}$. so that

$$c_p\left(h^*\right)\left(h_p - h_p^*\right) \geq u_{ij}\left(h_p - h_p^*\right) \tag{5.80}$$

including the case of $\left(h_p - h_p^*\right) < 0$, for then

$$h_p^* > h_p \geq 0 \Longrightarrow c_p(h^*) = u_{ij}$$

Therefore, from (5.80), upon summing over paths, we have at once the variational inequality (5.78).

(ii) [$VI\,(C, \Upsilon) \Longrightarrow UE\,(C, \Upsilon)$] The Kuhn-Tucker conditions for (5.78) are

$$\begin{aligned} c_p\left(h^*\right) - u_{ij} - \rho_p &= 0 \qquad \forall\,(i,j) \in \mathcal{W},\ p \in \mathcal{P}_{ij} \\ \rho_p h_p &= 0 \qquad \forall\,(i,j) \in \mathcal{W},\ p \in \mathcal{P}_{ij} \\ \rho_p &\geq 0 \qquad \forall\,(i,j) \in \mathcal{W},\ p \in \mathcal{P}_{ij} \end{aligned}$$

which are easily seen to yield the conditions that define $UE\,(C, \Upsilon)$. ■

5.3.7 *Existence and Uniqueness*

There is an existence and uniqueness theory for finite-dimensional games and variational inequalities. The relevant starting point for developing an existence theory for finite-dimensional nonextremal noncooperative games is, not surprisingly, Brouwer's existence theorem for fixed-point problems in $\Re^n$; we will employ the following version, slightly different from that presented in Chapter 4:

Theorem 5.8. *Brouwer's fixed-point theorem. If S is a convex, nonempty, and compact set, and $F(x)$ is continuous on S, then the fixed-point problem $FPP\,(F, \Lambda)$ has a solution.*

Proof. See Todd (1976). ■

Not surprisingly then, we also have the following existence result:

Theorem 5.9. *Stampacchia existence theorem. If Λ is convex, nonempty, and compact and $F(x)$ is continuous on Λ, then $VI\,(F, \Lambda)$ has a solution.*

Proof. By the given, $FPP_{\min}\,(F, \Lambda)$ satisfies the regularity conditions of Brouwer's fixed-point theorem and must, therefore, have a solution. It is immediate that $VI\,(F, \Lambda)$ also has a solution, since by Lemma 5.2 we know that any solution of $FPP_{\min}\,(F, \Lambda)$ is a solution of $VI\,(F, \Lambda)$. ■

This last theorem has an important implication for equilibria of Nash and Nash-like noncooperative games. In particular, we have:

Corollary 5.1. *Existence of $NE\,(\Theta, \Omega)$. Assume Ω is a convex, nonempty and compact set. A Nash equilibrium $NE\,(\Theta, \Omega)$ exists when $\theta_i(x^i, x^{-i})$ is convex and continuously differentiable with respect to x^i for all $x = \left(x^i, x^{-i}\right) \in \Omega$ and every $i \in [1, N]$.*

Proof. The result is immediate from Theorems 5.9 and 5.5. ■

We now discuss the uniqueness of solutions to variational inequality problems. To this end, we introduce the notion of monotonicity of a vector function:

Definition 5.11. *Montotonically increasing function. A function* $F(y) : \Re^n \longrightarrow \Re^n$ *is monotonically increasing on* Λ *if*

$$[F(y^1) - F(y^2)]^T (y^1 - y^2) \geq 0 \tag{5.81}$$

for all $y^1, y^2 \in \Lambda$.

We also introduce at this time the notion of strict monotonicity:

Definition 5.12. *Strictly monotonically increasing function. A function* $F(y) : \Re^n \longrightarrow \Re^n$ *is strictly monotonically increasing on* Λ *if*

$$[F(y^1) - F(y^2)]^T (y^1 - y^2) > 0 \tag{5.82}$$

for all $y^1, y^2 \in \Lambda$ *such that* $y^1 \neq y^2$.

Of course monotone decreasing versions of the above definitions are obtained by reversing the directions of the inequalities. The notion of strict monotonicity allows us to establish the following uniqueness result:

Theorem 5.10. *VI* (F, Λ) *uniqueness. If* $y \in \Lambda \subseteq \Re^n$ *is a solution of VI* (F, Λ) *and* $F(x)$ *is strictly monotonically increasing then* y *is unique.*

Proof. Suppose there are two solutions $y^1 \in \Lambda$ and $y^2 \in \Lambda$, where $y^1 \neq y^2$; as such the following variational inequalities obtain:

$$F(y^1)(y^2 - y^1) \geq 0 \text{ and } F(y^2)(y^1 - y^2) \geq 0 \tag{5.83}$$

Adding these inequalities leads to

$$[F(y^1) - F(y^2)]^T (y^1 - y^2) \leq 0, \tag{5.84}$$

which contradicts strict monotonicity (5.82). Hence $y^1 = y^2$, and any solution is unique. ■

There is an intimate relationship between differentiable convex functions and monotonically increasing functions that is important to the qualitative analysis of variational inequalities. That result is:

Theorem 5.11. *Relationship of convexity and monotonicity. If the differentiable function* $E(x) : \Lambda \subseteq \Re^n \longrightarrow \Re^n$ *is (strictly) convex for all* $x \in \Lambda$, *then its gradient* $\nabla E(x)$ *is (strictly) monotonically increasing for all* $x \in \Lambda$.

Proof. Convexity and differentiability of $E(x)$ ensure that

$$E\left(y^1\right) \geq E\left(y^2\right) + \left[\nabla E\left(y^2\right)\right]^T \left(y^1 - y^2\right) \tag{5.85}$$

$$E\left(y^2\right) \geq E\left(y^1\right) + \left[\nabla E\left(y^1\right)\right]^T \left(y^2 - y^1\right) \tag{5.86}$$

for all $y^1, y \in \Lambda$. Adding these two inequalities leads directly to

$$0 \geq [\nabla E(y_2)]^T \left(y^1 - y^2\right) + \left[\nabla E\left(y^1\right)\right]^T \left(y^2 - y^1\right) \tag{5.87}$$

which is easily manipulated to obtain

$$-[\nabla E(y_2)]^T \left(y^1 - y^2\right) + \left[\nabla E\left(y^1\right)\right]^T \left(y^1 - y^2\right) \geq 0 \tag{5.88}$$

or

$$\left\{\left[\nabla E\left(y^1\right)\right]^T - [\nabla E(y_2)]^T\right\} \left(y^1 - y^2\right) \geq 0 \tag{5.89}$$

which is recognized as the condition defining the monotonically increasing nature of $E(.)$. ■

5.4 Sensitivity Analysis of Variational Inequalities

For both extremal and nonextremal problems, we are frequently interested in estimating a new solution from a known solution after changes in model parameters. A very elegant theory that is also quite practical can be developed for such sensitivity analyses of variational inequalities. In light of our previous development, it should be clear to the reader that the theory of variational inequality sensitivity analysis that we are about to present is also relevant to Nash games.

We begin our discussion of sensitivity analysis with a reminder of what is meant by sensitivity analysis. In our presentation, sensitivity analysis is the analysis of the impact of parameter perturbations on the solutions of variational inequalities. We modify the statement of the variational inequality problem to include explicit reference to parameters that are determined external to the problem:

Definition 5.13. *Perturbed variational inequality. Given a vector of exogenous parameter perturbations $\xi \in \Re^s$, a function $h(x^*;\xi) : \Re^n \longrightarrow \Re^q$, a function $g(x^*;\xi) : \Re^n \longrightarrow \Re^m$, a function $F(x;\xi) : \Lambda \rightarrow \Re^n$ and the feasible set*

$$\Lambda(\xi) = \{x \in \Re^n : h(x,\xi) = 0, g(x,\xi) \leq 0\}, \tag{5.90}$$

the perturbed variational inequality problem VI $(F, \Lambda;\xi)$ is to find a vector y such that

$$\left.\begin{array}{l} y \in \Lambda(\xi) \\ [F(y;\xi)]^T (x - y) \geq 0 \quad \forall\, x \in \Lambda(\xi) \end{array}\right\} VI(F, \Lambda;\xi) \tag{5.91}$$

We are now able to state and prove the following preliminary result:

Theorem 5.12. *Continuous differentiable functions of perturbations. Let*

$$x^* \in \Lambda(0) = \{x \in \Re^n : h(x,0) = 0, G(x,0) \leq 0\}$$

be a solution of VI $(F, \Lambda; 0)$ *such that: (i) local sufficiency of any solution of VI* $(F, \Lambda; 0)$ *is assured; (ii) the gradients* $\nabla g_i\ (x^*, 0)$ *for all* i *such that* $g_i\ (x^*, 0) = 0$ *together with the gradients* $\nabla h_i\ (x^*, 0)$ *for all* $i \in [1, q]$ *are linearly independent; and (iii) the strict complementary slackness condition*

$$\pi_i > 0 \text{ when } g_i\left(x^*, 0\right) = 0 \tag{5.92}$$

is satisfied. Then the multipliers μ^* *and* π^* *associated with* x^* *are unique. Also, in a neighborhood of* $\xi = 0$*, there exists a unique once continuously differentiable function*

$$\begin{bmatrix} x\ (\xi) \\ \mu\ (\xi) \\ \pi\ (\xi) \end{bmatrix} \tag{5.93}$$

where $x\ (\xi)$ *is a locally unique solution of VI* $(F, \Lambda; \xi)$ *and* $\mu\ (\xi)$ *and* $\pi\ (\xi)$ *are unique associated Kuhn-Tucker multipliers. Furthermore, in a neighborhood of* $\xi = 0$*, the gradients of constraints binding at* $x\ (\xi)$ *are linearly independent.*

Proof. For $\xi = 0$ the solutions of the perturbed and unperturbed variational inequalities are of course identical:

$$\begin{bmatrix} x^* \\ \mu^* \\ \pi^* \end{bmatrix} = \begin{bmatrix} x\ (0) \\ \mu\ (0) \\ \pi\ (0) \end{bmatrix} \tag{5.94}$$

From Theorem 5.2 we have

$$F(x, \xi) + [\nabla g(x, \xi)]^T \pi + [\nabla h(x, \xi)]^T\ \mu = 0 \tag{5.95}$$

$$\pi^T g\ (x, \xi) = 0 \tag{5.96}$$

$$h\ (x, \xi) = 0 \tag{5.97}$$

$$\pi \geq 0 \tag{5.98}$$

A system of equations with the structure of (5.95) through (5.98) has a Jacobian with respect to x, μ, and π that is nonsingular at π^*, μ^*, and π^* when $\xi = 0$. Consequently the implicit function theorem may be invoked and the conclusions of the present theorem follow immediately. ■

The preceding theorem ensures that the perturbed variational inequality has well-behaved primal and dual solutions that vary continuously with the perturbations and can be differentiated with respect to those perturbations.

It is now rather easy to establish our central result on sensitivity analysis of variational inequalities:

Corollary 5.2. *First-order approximate solutions of VI* $(F, \Lambda; \xi)$ *near* $\xi = 0$*. Under the assumptions of Theorem 5.12, a first-order approximation of the solution to VI* $(F, \Lambda; \xi)$ *is*

$$y(\xi) = \begin{bmatrix} x(\xi) \\ \mu(\xi) \\ \pi(\xi) \end{bmatrix} = \begin{bmatrix} x^* \\ \mu^* \\ \pi^* \end{bmatrix} + \left[J_y(0)\right]^{-1} \left[-J_\xi^*(0)\right] \xi \tag{5.99}$$

where $J_y(\xi)$ is the Jacobian matrix of the system (5.95), (5.96), (5.97), and (5.98) with respect to $y(\xi)$ evaluated at $(y(\xi), \xi)$ and $J_\xi(\xi)$ is the Jacobian of the same system with respect to ξ also evaluated at $(y(\xi), \xi)$.

Proof. The proof is given by Tobin (1986) and parallels an earlier result by Fiacco and McCormick (1990) for perturbed mathematical programs. ■

5.5 The Diagonalization Algorithm

Since Nash and many Nash-like equilibria may be articulated as variational inequalities, we only focus here in this section on algorithms for variational inequalities. On the surface, it would appear that variational inequalities can be solved by reformulating them as mathematical programs using Lemma 5.3. However, that result depends on the introduction of an objective function that involves a line integral. Line integrals are not generally single valued; in fact their value depends on the path of integration one employs. As this is a somewhat subtle point, an example is warranted. Consider the line integral

$$I = \oint_{(a_1,a_2)}^{(b_1,b_2)} [F(x)]^T dx \tag{5.100}$$

for which $x \in \Re^2$ and $F : \Re^2 \longrightarrow \Re^2$. In summation notation we write

$$I = \sum_{i=1}^{2} \int_{a_i}^{b_i} F_i(x_1, x_2)\, dx_i \tag{5.101}$$

Let

$$\begin{aligned} F_1(x_1, x_2) &= x_1 + 2x_2 \\ F_2(x_1, x_2) &= x_1 + x_2 \\ a_1 &= a_2 = 0 \\ b_1 &= b_2 = 1 \end{aligned}$$

and consider two distinct paths of integration:

Path	1st segment	2nd segment
1	$x_1 = 0, x_2 \in [0,1]$	$x_1 \in [0,1], x_2 = 1$
2	$x_1 \in [0,1], x_2 = 0$	$x_1 = 1, x_2 \in [0,1]$

For path 1 we have

$$I = \int_0^1 F_2\,(0, x_2)\,dx_2 + \int_0^1 F_1\,(x_1, 1)\,dx_1$$
$$= \int_0^1 x_2 dx_2 + \int_0^1 (x_1 + 2)\,dx_1 = 3 \tag{5.102}$$

For path 2 we have

$$I = \int_0^1 F_1\,(x_1, 0)\,dx_1 + \int_0^1 F_2\,(1, x_2)\,dx_2$$
$$= \int_0^1 x_1 dx_1 + \int_0^1 (1 + x_2)\,dx_2 = 2 \tag{5.103}$$

Evidently the value of this line integral depends on the path of integration.

In fact, it is well known that a line integral

$$I = \oint_a^b F(x)dx = \sum_{i=1}^{n} \int_{a_i}^{b_i} F_i\,(x_1, x_2)\,dx_i \tag{5.104}$$

where $a, b, x \in \Re^n$ and $F(x) : \Re^n \longrightarrow \Re^n$ has a value independent of the path of integration if and only if

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i} \quad \forall i, j \in [1, n] \tag{5.105}$$

The restrictions (5.105) are known as symmetry conditions since they make the Jacobian matrix

$$J(F) \equiv \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_n} \end{pmatrix} \tag{5.106}$$

symmetric. It is significant that one class of functions $F(x) : \Re^n \longrightarrow \Re^n$ always leads to a symmetric $J(F)$ and thereby satisfaction of (5.105); that is the class of functions known as separable functions for which each scalar component has only an own-variable dependence, which we express symbolically as

$$F_i = F_i\,(x_i) \quad \forall i \in [1, n] \tag{5.107}$$

By inspection we see that the Jacobian matrix for the vector function $F(x)$ whose scalar components obey (5.107) is a diagonal and therefore symmetric matrix.

5.5.1 The Algorithm

The algorithm we emphasize in this section is called the *diagonalization algorithm* or *diagonalization* for short; it is an algorithmic philosophy very similar to the Gauss-Seidel method[1] familiar from the numerical analysis literature. Diagonalization is appealing for solving finite-dimensional variational inequalities because the resulting subproblems are all nonlinear programs that can be efficiently solved with well-understood nonlinear programming algorithms, which are often available in the form of commercial software. This fact not withstanding, diagonalization may fail to converge and its use on large-scale problems can be frustrating.

The diagonalization algorithm rests on the creation of separable functions at each iteration k of the form

$$F_i^k(x_i) \equiv F_i\left(x_i, x_j = x_j^k \ \ \forall\ j \neq i\right) \tag{5.108}$$

Evidently the functions $F_i^k(x_i)$ are separable by construction, so that the Jacobian of $F^k = (\ldots, F_i^k, \ldots)^T$ is diagonal; hence, the name of the method. The diagonalization algorithm may be stated as follows:

Diagonalization Algorithm for VI(F, Λ)

Step 0. Initialization. Determine an initial feasible solution $x^0 \in \Lambda$ and set $k = 0$.

Step 1. Solve diagonalized variational inequality. Form the separable functions $F_i^k(x_i)$ for all $i \in [1, n]$ and solve the associated diagonalized variational inequality problem. That is, find $x^{k+1} \in \Lambda$ such that

$$\sum_{i=1}^{n} F_i^k(x_i^{k+1})(x_i - x_i^{k+1}) \geq 0 \quad \forall x \in \Lambda \tag{5.109}$$

Step 2. Stopping test and updating. For $\zeta \in \Re^1_{++}$, a preset tolerance, if

$$\max_{i \in [1,n]} |x_i^{k+1} - x_i^k| < \zeta$$

stop; otherwise set $k = k + 1$ and go to Step 1.

Note that the variational inequalities of Step 1 of the above algorithm may be solved using the nonlinear program

$$\min J(x) = \sum_i \int_0^{x_i} F_i^k(z_i)\, dz_i \quad \text{s.t. } x \in \Lambda\,, \tag{5.110}$$

[1] See Ortega and Rheinboldt (2000) for a typology of iterative algorithms in numerical analysis.

where the z_i are dummy variables of integration, because the conditions needed to invoke Lemma 5.3 are in force since the integral in (5.110) is an ordinary integral, not a line integral, and no symmetry restrictions need be imposed because the functions $F_i^k(\cdot)$ are separable. Thus, the diagonalization algorithm involves, in effect, the solution of a sequence of separable problems, each of which may be expressed as a well-defined mathematical program.

5.5.2 *Converence of Diagonalization*

To state a convergence result for the diagonalization algorithm, we will use the G-norm of a vector x, which is defined by

$$\|x\|_G \equiv \left(x^T G x\right)^{1/2} \tag{5.111}$$

Referring back to the variational inequality problem $VI(F, \Gamma)$ and letting D and B denote respectively the diagonal and off-diagonal portions of $[\nabla F(x^*)]^T \equiv J[F(x^*)]$ (the Jacobian of $F(x)$ evaluated at $x = x^*$), we consider the following theorem due to Pang and Chan (1982):

Theorem 5.13. *Convergence of diagonalization. Let D and B denote, respectively, the diagonal and off-diagonal portions of $\nabla f(x^*)$. If x^* is a solution of $VI(F, \Gamma)$ and*

(1) Γ *is convex*

(2) $F(x)$ *is differentiable* $\forall x \in \Gamma$

(3) $F(x)$ *is continuously differentiable in a neighborhood of* x^*

(4) $\dfrac{\partial F_i(x)}{\partial x_i} \geq 0 \quad \forall\, i \in [1, n],\ x \in \Gamma$

(5) $\dfrac{\partial F_i(x^*)}{\partial x_i} > 0 \quad \forall\, i \in [1, n]$

(6) $\| D^{-1/2} B D^{-1/2} \| < 1,$

then, provided that the initial vector x^0 is chosen in a suitable neighborhood of x^*, the diagonalization algorithm will converge to x^*.

Proof. See Pang and Chan (1982). ■

It is interesting to note that both Pang and Chan (1982) and Dafermos (1983) give global proofs of convergence of the diagonalization method by invoking more restrictive regularity conditions than those employed above. Also, diagonalization and the family of related iterative algorithms which Pang and Chan (1982) and Dafermos (1983) discuss may converge under circumstances that do not fulfill the known convergence theory. Nonetheless, examples of nonconvergence are known and the method must be used with great caution.

5.5.3 A Nonnetwork Example of Diagonalization

In this section we consider an example given originally by Tobin (1986). In particular, we study the following variational inequality: find $\left(x_1^*, x_2^*\right)^T \in \Gamma$ such that

$$F_1\left(x_1^*, x_2^*\right)\left(x_1 - x_1^*\right) + F_2\left(x_2^*, x_2^*\right)\left(x_2 - x_2^*\right) \geq 0 \quad \forall \left(x_1, x_2\right)^T \in \Gamma \quad (5.112)$$

where

$$\Gamma = \left\{\left(x_1, x_2\right)^T : g_1\left(x_1, x_2\right) \leq 0, g_2\left(x_1, x_2\right) \leq 0, g_3\left(x_1, x_2\right) \leq 0\right\}$$

and

$$\begin{aligned} F_1\left(x_1, x_2\right) &= x_1 - 5 \\ F_2\left(x_1, x_2\right) &= .1x_1x_2 + x_2 - 5 \\ g_1\left(x_1, x_2\right) &= -x_1 \leq 0 \quad (\lambda_1) \\ g_2\left(x_1, x_2\right) &= -x_2 \leq 0 \quad (\lambda_2) \\ g_3\left(x_1, x_2\right) &= x_1 + x_2 - 1 \leq 0 \quad (\lambda_3) \end{aligned}$$

By inspection, the Jacobian

$$\nabla F\left(x_1, x_2\right) = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_2}{\partial x_1} \\ \dfrac{\partial F_1}{\partial x_2} & \dfrac{\partial F_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 1 & .1x_2 \\ 0 & .1x_1 + 1 \end{pmatrix} \quad (5.113)$$

is asymmetric and, consequently, we cannot directly construct an equivalent optimization problem with a single valued objective function. In particular, we note that an equivalent optimization problem must have as an objective the line integral

$$Z = \oint_{(0,0)}^{(x_1,x_2)} \left[F\left(z\right)\right]^T dz \quad (5.114)$$

We diagonalize by constructing

$$F_1^k\left(x_1\right) \equiv F_1\left(x_1, x_2^k\right) = x_1 - 5 \quad (5.115)$$

$$F_2^k\left(x_2\right) \equiv F_2\left(x_1^k, x_2\right) = \left(.1x_1^k + 1\right)x_2 - 5 \quad (5.116)$$

where $\left(x_1^k, x_2^k\right)^T \in \Gamma$ is the current approximate solution. Thus, the mathematical program solved to find x^{k+1} has the objective

$$\begin{aligned}\min Z^k &\equiv \int_0^{x_1} F_1^k(z_1)\,dz_1 + \int_0^{x_2} F_2^k(z_2)\,dz_2 \\ &= \int_0^{x_1} (z_1 - 5)\,dz_1 + \int_0^{x_2} \left[\left(.1x_1^k + 1\right) z_2 - 5\right] dz_2 \\ &= \frac{1}{2}(x_1)^2 - 5x_1 + \frac{1}{2}\left(.1x_1^k + 1\right)(x_2)^2 - 5x_2 \quad \text{s.t. } (x_1, x_2)^T \in \Gamma\end{aligned} \tag{5.117}$$

We denote the solution of (5.117) by $\left(x_1^{k+1}, x_2^{k+1}\right)^T$, which of course satisfies the Kuhn-Tucker conditions:

$$\frac{\partial Z^k}{\partial x_1} + \lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1} + \lambda_3 \frac{\partial g_3}{\partial x_1} = x_1 - 5 - \lambda_1 + \lambda_3 = 0 \tag{5.118}$$

$$\frac{\partial Z^k}{\partial x_2} + \lambda_1 \frac{\partial g_1}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2} + \lambda_3 \frac{\partial g_3}{\partial x_2} = \left(.1x_1^k + 1\right) x_2 - 5 - \lambda_2 + \lambda_3 = 0 \tag{5.119}$$

$$\lambda_1 x_1 = 0 \tag{5.120}$$

$$\lambda_2 x_2 = 0 \tag{5.121}$$

$$\lambda_3 (x_1 + x_2 - 1) = 0 \tag{5.122}$$

$$\lambda_1, \lambda_2, \lambda_3 \geq 0 \tag{5.123}$$

Assuming a primal solution in the first quadrant and that $g_3(x_1, x_2)$ binds at optimality, as is easily verified by graphical solution of (5.117), these conditions reduce to

$$\begin{aligned} x_1 + \lambda_3 &= 5 \\ \left(.1x_1^k + 1\right) x_2 + \lambda_3 &= 5 \\ x_1 + x_2 &= 1 \\ \lambda_1 = \lambda_2 &= 0 \end{aligned}$$

which yield the convenient formulae

$$x_1 = \frac{x_1^k + 10}{x_1^k + 20} \tag{5.124}$$

$$x_2 = \frac{10}{x_1^k + 20} \tag{5.125}$$

$$\lambda_1 = \lambda_2 = 0 \tag{5.126}$$

$$\lambda_3 = 2\frac{2x_1^k + 45}{x_1^k + 20} \tag{5.127}$$

It must be noted that generally the mathematical programming subproblems resulting from diagonalization are not this simple and one must select an appropriate numerical algorithm for their solution.

With the results developed above, we can describe the iterations of the diagonalization algorithm:

Step 0. Initialization. Pick $\left(x_1^0, x_2^0\right)^T = (1,0)^T$ and set $k = 0$.

$k = 0$, Step1. Solve the diagonalized variational inequality using (5.124) to find:

$$\begin{pmatrix} x_1^1 \\ x_2^1 \end{pmatrix} = \begin{pmatrix} \dfrac{x_1^0 + 10}{x_1^0 + 20} \\ \dfrac{10}{x_1^0 + 20} \end{pmatrix} = \begin{pmatrix} \dfrac{11}{21} \\ \dfrac{10}{21} \end{pmatrix} = \begin{pmatrix} 0.52381 \\ 0.47619 \end{pmatrix}$$

$k = 0$, Step 2. Updating. Set $k = 0 + 1 = 1$.

$k = 1$, Step 1. Solve the diagonalized variational inequality using (5.124) to find:

$$\begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix} = \begin{pmatrix} \dfrac{x_1^1 + 10}{x_1^1 + 20} \\ \dfrac{10}{x_1^1 + 20} \end{pmatrix} = \begin{pmatrix} \dfrac{0.52381 + 10}{0.52381 + 20} \\ \dfrac{10}{0.52381 + 20} \end{pmatrix} = \begin{pmatrix} 0.51276 \\ 0.48724 \end{pmatrix}$$

$k = 1$, Step 2. Updating. Set $k = 1 + 1 = 2$.

$k = 2$, Step 1. Solve the diagonalized variational inequality using (5.124) to find:

$$\begin{pmatrix} x_1^3 \\ x_2^3 \end{pmatrix} = \begin{pmatrix} \dfrac{x_1^2 + 10}{x_1^2 + 20} \\ \dfrac{10}{x_1^2 + 20} \end{pmatrix} = \begin{pmatrix} \dfrac{0.51276 + 10}{0.51276 + 20} \\ \dfrac{10}{0.51276 + 20} \end{pmatrix} = \begin{pmatrix} 0.5125 \\ 0.4875 \end{pmatrix}$$

$k = 2$, Step 2. Stopping. Assuming a stopping tolerance $\varepsilon = .001$, we see that

$$\begin{aligned} \max_{i \in [1,2]} |x_i^3 - x_i^2| &= \max \{|0.51276 - 0.5125|, |0.48724 - 0.4875|\} \\ &= .00026 < .001 \\ \Longrightarrow \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} &\cong \begin{pmatrix} x_1^3 \\ x_2^3 \end{pmatrix} = \begin{pmatrix} 0.5125 \\ 0.4875 \end{pmatrix} \end{aligned} \tag{5.128}$$

Note also that the dual variables associated with solution (5.128) are

$$\lambda_1^* = \lambda_2^* = 0 \tag{5.129}$$

$$\lambda_3^* = 2\frac{2x_1^* + 45}{x_1^* + 20} = 2\frac{2\,(0.5125) + 45}{(0.5125) + 20} = 4.4875 \tag{5.130}$$

By inspection, the inequality constraint functions are linearly independent and convex. Consequently, the Kuhn-Tucker conditions for this variational inequality are necessary and sufficient, so that any solution to

$$F_j\left(x_1^*, x_2^*\right) + \sum_{i=1}^{3} \lambda_i \frac{\partial g_i\left(x_1^*, x_2^*\right)}{\partial x_i} = 0 \quad j = 1, 2 \tag{5.131}$$

$$\lambda_i g_i\left(x_1^*, x_2^*\right) = 0 \quad i = 1, 2, 3 \tag{5.132}$$

$$\lambda_i \geq 0 \quad i = 1, 2, 3 \tag{5.133}$$

is the desired global solution. These conditions yield

$$x_1^* - 5 - \lambda_1^* + \lambda_3^* = 0 \tag{5.134}$$

$$.1x_1^* x_2^* + x_2^* - 5 - \lambda_2^* + \lambda_3^* = 0 \tag{5.135}$$

$$\lambda_1^* x_1^* = 0 \tag{5.136}$$

$$\lambda_2^* x_2^* = 0 \tag{5.137}$$

$$\lambda_3^*\left(x_1^* + x_2^* - 1\right) = 0 \tag{5.138}$$

$$\lambda_1^*, \lambda_2^*, \lambda_3^* \geq 0 \tag{5.139}$$

which are subtly different than conditions (5.118) through (5.123). In fact, (5.136), (5.137), (5.138), and (5.139) are seen by inspection to be satisfied, while (5.134) and (5.135) give

$$x_1^* - 5 - \lambda_1^* + \lambda_3^* = 0.5125 - 5 - 0 + 4.4875 = 0 \tag{5.140}$$

$$\begin{aligned} .1x_1^* x_2^* + x_2^* - 5 - \lambda_2^* + \lambda_3^* &= .1\,(0.5125)\,(0.4875) + 0.4875 - 5 - 0 + 4.4875 \\ &= -1.5625 \times 10^{-5} \end{aligned} \tag{5.141}$$

That is, the variational-inequality Kuhn-Tucker conditions are approximately satisfied by our solution obtained from the diagonalization algorithm.

Now imagine that the function $F_2\,(x_1, x_2) = .1x_1x_2 + x_2 - 5$ is perturbed according to

$$F_2\,(x_1, x_2; \xi) = (.1 + \xi)\,x_1x_2 + x_2 - 5 \tag{5.142}$$

where now $\xi \in \Re^1$. To apply the sensitivity analysis results derived in Section 5.4, we employ the Kuhn-Tucker system for the perturbed problem:

$$x_1 - 5 - \lambda_1 + \lambda_3 = 0 \tag{5.143}$$
$$(.1 + \xi)\, x_1 x_2 + x_2 - 5 - \lambda_2 + \lambda_3 = 0 \tag{5.144}$$
$$\lambda_1 x_1 = 0 \tag{5.145}$$
$$\lambda_2 x_2 = 0 \tag{5.146}$$
$$\lambda_3 (x_1 + x_2 - 1) = 0 \tag{5.147}$$

where of course $\lambda_1, \lambda_2, \lambda_3 \geq 0$. The relevant Jacobians of this system are

$$J_y(\xi) = \begin{pmatrix} 1 & 0 & -1 & 0 & 1 \\ (.1+\xi)\, x_2 & (.1+\xi)\, x_1 & 0 & -1 & 1 \\ \lambda_1 & 0 & x_1 & 0 & 0 \\ 0 & \lambda_2 & 0 & x_2 & 0 \\ \lambda_3 & \lambda_3 & 0 & 0 & x_1 + x_2 - 1 \end{pmatrix}$$

$$J_\xi^*(\xi) = \begin{pmatrix} 0 \\ x_1 x_2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Consequently

$$\begin{bmatrix} x_1(\xi) \\ x_2(\xi) \\ \lambda_1(\xi) \\ \lambda_2(\xi) \\ \lambda_3(\xi) \end{bmatrix} = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \lambda_1(0) \\ \lambda_2(0) \\ \lambda_3(0) \end{bmatrix} + \left[J_y(0)\right]^{-1} \left[-J_\xi^*(0)\right] \xi$$

$$= \begin{pmatrix} x_1^* \\ x_2^* \\ \lambda_1^* \\ \lambda_2^* \\ \lambda_3^* \end{pmatrix} - \begin{pmatrix} 1 & 0 & -1 & 0 & 1 \\ (.1+\xi)\, x_2^* & (.1+\xi)\, x_1^* & 0 & -1 & 1 \\ \lambda_1^* & 0 & x_1^* & 0 & 0 \\ 0 & \lambda_2^* & 0 & x_2^* & 0 \\ \lambda_3^* & \lambda_3^* & 0 & 0 & x_1^* + x_2^* - 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ x_1^* x_2^* \\ 0 \\ 0 \\ 0 \end{pmatrix} \xi$$

$$= \begin{pmatrix} 0.5125 + 1.4006 \times 10^9 \dfrac{\xi}{5.6199 \times 10^9 + 1.4015 \times 10^8 \xi} \\ 0.4875 - 1.4006 \times 10^9 \dfrac{\xi}{5.6199 \times 10^9 + 1.4015 \times 10^8 \xi} \\ 0 \\ 0 \\ 4.4875 - 1.4006 \times 10^9 \dfrac{\xi}{5.6199 \times 10^9 + 1.4015 \times 10^8 \xi} \end{pmatrix} \tag{5.148}$$

Use of (5.148) leads to the following table of results:

perturbation	exact	1^{st} order	exact	1^{st} order	exact	1^{st} order
ξ	$x_1^*(\xi)$	$x_1(\xi)$	$x_2^*(\xi)$	$x_2(\xi)$	$\lambda_3^*(\xi)$	$\lambda_3(\xi)$
0	0.5125	0.5125	0.4875	0.4875	4.4875	4.4875
.01	0.5137	0.5150	0.4863	0.4850	4.4863	4.4850
.05	0.5187	0.5250	0.4813	0.4751	4.4813	4.4751
.1	0.5249	0.5374	0.4751	0.4626	4.4751	4.4626
.15	0.5311	0.5497	0.4689	0.4503	4.4689	4.4503

Note that in the above table we have used the notation $x_1^*(\xi)$, $x_2^*(\xi)$, and $\lambda_3^*(\xi)$ to denote the exact solutions of the perturbed problem and the notation $x_1(\xi)$, $x_2(\xi)$, and $\lambda_3(\xi)$ to denote the first-order approximate solutions found using sensitivity analysis. The exact solutions were found using separate software and are included for comparison. Even for rather large pertubations, the first-order approximations stemming from the sensitivity analysis are quite good. This accuracy for large perturbations, although not guaranteed, is not uncommon for the sensitivity analysis theory we have presented.

5.6 Gap Function Methods for *VI* (F, Λ)

There is a special class of functions associated with variational inequality problems, so-called gap functions, which forms the foundation of a family of algorithms that are sometimes very effective for solving $VI(F, \Lambda)$. A gap function has two important and advantageous properties: (1) it is always nonnegative and (2) it has zero value if and only if we have a solution of the corresponding variational inequality.

5.6.1 Gap Function Defined

Formally, we define a gap function for $VI(F, \Lambda)$ as follows:

Definition 5.14. *Gap function. A function* $\zeta : \Lambda \subseteq \Re^n \longrightarrow \Re_+$ *is called a gap function for* $VI(F, \Lambda)$ *when the following statements hold:*

1. $\zeta(y) \geq 0$ *for all* $y \in \Lambda$
2. $\zeta(y) = 0$ *if and only if* y *is the solution of* $VI(F, \Lambda)$

Clearly, a gap function with the properties of Definition 5.14 allows us to reformulate $VI(F, \Lambda)$ as an optimization problem, namely as

$$\min_{y \in \Lambda} \zeta(y) \tag{5.149}$$

An optimal solution of (5.149) solves $VI(F, \Lambda)$ provided $\zeta(y)$ may be driven to zero.

5.6.2 The Auslender Gap Function

A gap function given by Auslender (1976) is the subject of the following result that employs the notation

$$\langle v, w \rangle \equiv v^T w$$

when $v, w \in \Re^n$:

Theorem 5.14. *Auslender's function is a gap function. The function*

$$\zeta\,(y) = \max_{x \in \Lambda} \langle F\,(y)\,, y - x \rangle \tag{5.150}$$

is a gap function for $VI(F, \Lambda)$, where Λ is convex.

Proof. The proof is in two parts:

(i) $[\zeta\,(y) \geq 0]$ To establish Property 1 of Definition 5.14, we observe that, when Λ is convex, a necessary and sufficient condition for $x \in \Lambda$ to solve (5.150) is

$$[F\,(y)]^T\,(z - x) \geq 0 \qquad \forall z \in \Lambda \tag{5.151}$$

Picking $z = y$, it is immediate that $\zeta\,(y)$ is a nonnegative function of y.

(ii) $[\zeta(y) = 0 \Longleftrightarrow VI(F, \Lambda]$ If $y \in \Lambda$ solves $VI(F, \Lambda)$ then

$$[F(y)]^T\,(x - y) \geq 0 \qquad \forall x \in \Lambda$$

or

$$[F(y)]^T\,(y - x) \leq 0 \qquad \forall x \in \Lambda \tag{5.152}$$

Comparing (5.152) to (5.150) assures $\zeta\,(y) = 0$, as required. To show $\zeta(y) = 0$ assures y solves $VI(F, \Lambda)$, let us assume it does not; that is, there exists $x \in \Lambda$ such that

$$[F(y)]^T\,(y - x) > 0 \tag{5.153}$$

However, (5.153) means that $\zeta\,(y) = 0$ cannot be the result of solving (5.150), which is a contradiction; therefore $\zeta\,(y) = 0$ assures y solves $VI(F, \Lambda)$. ■

For the Auslender gap function, we may rewrite (5.149) as

$$\min_{y \in \Lambda} \zeta\,(y) = \min_{y \in \Lambda} \max_{x \in \Lambda} \langle F\,(y)\,, y - x \rangle = \min_{y \in \Lambda} \left\{ \langle F\,(y)\,, y \rangle + \max_{x \in \Lambda} \langle F\,(y)\,, -x \rangle \right\} \tag{5.154}$$

a format that reveals the underlying min-max nature of the gap function perspective for solving variational inequalities. Note further that the Auslender gap function $\zeta\,(y)$ is not in general differentiable, even if F is differentiable.

5.6.3 Fukushima-Auchmuty Gap Functions

Auchmuty (1989) and Fukushima (1992) independently suggested a class of differentiable gap functions of the form

$$\zeta_\alpha(y) = \max_{x \in \Lambda} \left\{ \langle F(y), y - x \rangle - \frac{\alpha}{2} \|y - x\|^2 \right\} \tag{5.155}$$

for $\alpha \in \Re^1_{++}$. Function (5.155) is differentiable whenever F is differentiable. In particular, recognizing that

$$\|y - x\|^2 = (y - x)^T (y - x) \equiv (y - x)^2 \ ,$$

its gradient with respect to y is given by

$$\nabla \zeta_\alpha(y) = F(y) + \langle \nabla F(y), y - x_\alpha \rangle - \alpha (y - x_\alpha)$$

where x_α denotes the unique maximizer of (5.155). The differentiability of (5.155) is due to the uniqueness and realized finiteness of x_α, which occurs because the objective function on the right-hand side of (5.155) is strongly convex in x.

Wu et al. (1993) proposed the following generalization of the Fukushima-Auchmuty gap function (5.155):

$$\zeta_\alpha(y) = \max_{x \in \Lambda} \{ \langle F(y), y - x \rangle - \alpha \phi(y, x) \} \tag{5.156}$$

In (5.156) ϕ is a function that satisfies the following conditions:

1. ϕ is continuously differentiable on $\Re^{2n}$;
2. ϕ is nonnegative on $\Re^{2n}$;
3. $\phi(y, \cdot)$ is strongly convex for any $y \in \Re^n$; and
4. $\phi(y, x) = 0$ if and only if $y = x$.

If (5.156) is a gap function, it gives rise to the following constrained mathematical program

$$\min_{y \in \Lambda} \zeta_\alpha(y)$$

which is equivalent to $VI(F, \Lambda)$ provided $\zeta_\alpha(y)$ may be driven to zero. That is, we are now ready to state and prove the following result:

Theorem 5.15. *The extended Fukushima-Auchmuty function is a gap function. The function (5.156) is a gap function for $VI(F, \Lambda)$, where Λ is convex.*

Proof. The proof is in two parts:

(i) [$\zeta_\alpha(y) \geq 0$] To establish Property 1 of Definition 5.14, we observe that (5.156) is equivalent to

$$\min_{x \in \Lambda} [F(y)]^T x + \alpha \phi(y, x) \tag{5.157}$$

Strong convexity of $\phi\,(y, \cdot)$ assures that the maximum in (5.156) and the minimum in (5.157) are bounded away from zero and may actually be attained. Furthermore, (5.157) has the necessary and sufficient condition

$$[F\,(y) + \alpha \nabla_x \phi\,(y, x)]^T\,(z - x) \geq 0 \qquad \forall z \in \Lambda$$

which upon picking $z = y$ becomes

$$[F\,(y)]^T\,(y - x) + \alpha\,[\nabla_x \phi\,(y, x)]^T\,(y - x) \geq 0 \qquad \forall z \in \Lambda \tag{5.158}$$

Because $\phi\,(y, \cdot)$ is strongly convex it is also convex so that

$$\phi\,(y, z) \geq \phi\,(y, x) + [\nabla_x \phi\,(y, x)]^T\,(y - x) \quad \forall z \in \Lambda$$

Taking $z = y$ in the last expression and noting that by the given $\phi\,(y, y) = 0$, we have

$$-\phi\,(y, x) \geq [\nabla_x \phi\,(y, x)]^T\,(y - x) \tag{5.159}$$

It is immediate from (5.158) and (5.159) that

$$\zeta_\alpha\,(y) = [F\,(y)]^T\,(y - x) - \alpha \phi\,(y, x) \geq 0 \qquad \forall z \in \Lambda \tag{5.160}$$

(ii) $[\zeta_\alpha(y) = 0 \Longleftrightarrow VI(F, \Lambda]$ If $y \in \Lambda$ solves $VI(F, \Lambda)$ then

$$[F(y)]^T\,(x - y) \geq 0 \qquad \forall x \in \Lambda$$

or

$$[F(y)]^T\,(y - x) \leq 0 \qquad \forall x \in \Lambda \tag{5.161}$$

So because $\phi\,(y, x) \geq 0$ we have

$$[F(y)]^T\,(y - x) - \alpha \phi\,(y, x) \leq 0 \qquad \forall x \in \Lambda \tag{5.162}$$

Comparing (5.162) and (5.160) assures $\zeta_\alpha\,(y) = 0$. On the other hand if $\zeta_\alpha(y) = 0$, then

$$[F(y)]^T\,(y - x) \geq [F(y)]^T\,(y - x) - \alpha \phi\,(y, x) = 0$$

which assures y solves $VI(F, \Lambda)$. ■

5.6.4 The D-Gap Function

The gap functions introduced above all lead to equivalent constrained mathematical programs. It is reasonable to ask whether there is a gap function that leads to an equivalent unconstrained mathematical program. In fact, the so-called *D-gap*

function proposed by Peng (1997) and generalized by Yamashita et al. (1997) is such a function. A D-gap function is the difference between two gap functions. The D-gap function we will consider is

$$\begin{aligned}\psi_{\alpha\beta}(y) &= \zeta_\alpha(y) - \zeta_\beta(y) \\ &= \max_{x\in\Lambda}\{\langle F(y), y-x\rangle - \alpha\phi(y,x)\} - \max_{x\in\Lambda}\{\langle F(y), y-x\rangle - \beta\phi(y,x)\}\end{aligned} \tag{5.163}$$

where $0 < \alpha < \beta$ and the conditions imposed on the function $\phi(y,x)$ are the same as those given in Section 5.6.3. The corresponding unconstrained mathematical program equivalent to $VI(F,\Lambda)$ is

$$\min_y \psi_{\alpha\beta}(y) \tag{5.164}$$

Moreover, the gradient of $\psi_{\alpha\beta}(y)$ is well defined. To express the gradient let us define $x_\alpha(y)$ and $x_\beta(y)$ such that

$$\max_{x\in\Lambda}\{\langle F(y), y-x\rangle - \alpha\phi(y,x)\} = \langle F(y), y - x_\alpha(y)\rangle - \alpha\phi(y, x_\alpha(y))$$

$$\max_{x\in\Lambda}\{\langle F(y), y-x\rangle - \beta\phi(y,x)\} = \langle F(y), y - x_\beta(y)\rangle - \beta\phi(y, x_\beta(y))$$

That is,

$$x_\alpha(y) = \arg\max_{x\in\Lambda}\{\langle F(y), y-x\rangle - \alpha\phi(y,x)\}$$

$$x_\beta(y) = \arg\max_{x\in\Lambda}\{\langle F(y), y-x\rangle - \beta\phi(y,x)\}$$

As a consequence we may rewrite (5.163) as

$$\begin{aligned}\psi_{\alpha\beta}(y) &= \langle F(y), y - x_\alpha(y)\rangle - \alpha\phi(y, x_\alpha(y)) - \langle F(y), y - x_\beta(y)\rangle \\ &\quad + \beta\phi(y, x_\beta(y)) \\ &= \langle F(y), x_\beta(y) - x_\alpha(y)\rangle + \beta\phi(y, x_\beta(y)) - \alpha\phi(y, x_\alpha(y))\end{aligned} \tag{5.165}$$

Since $\phi(y,\cdot)$ is strongly convex in y and Λ is convex and compact, $x_\alpha(y)$ and $x_\beta(y)$ are unique and realized as vectors whose components are finite. From (5.165) the gradient of $\psi_{\alpha\beta}(y)$ is readily seen to be

$$\nabla\psi_{\alpha\beta}(y) = \nabla F(y)\left(x_\beta(y) - x_\alpha(y)\right) + \beta\nabla_y\phi(y, x_\beta(y)) - \alpha\nabla_y\phi(y, x_\alpha(y))$$

For detailed proofs of the assertions we have made concerning $\psi_{\alpha\beta}(y)$ see Yamashita et al. (1997).

5.6.5 Gap Function Numerical Example

Once a differentiable gap function has been formed for $VI(F, \Lambda)$, it is used to create a nonlinear program that may be solved by conventional nonlinear programming methods. This is now illustrated for the D-gap function:

D-Gap Algorithm for $VI(F, \Lambda)$

Step 0. Initialization. Initialization. Determine an initial feasible solution $y^0 \in \Re^n$ and set $k = 0$.

Step 1. Finding the steepest descent direction. Find the gradient of the D-gap function:

$$\nabla \psi_{\alpha\beta}\left(y^k\right) = \nabla F\left(y^k\right)\left(x_\beta\left(y^k\right) - x_\alpha\left(y^k\right)\right) + \beta \nabla_y \phi\left(y^k, x_\beta\left(y^k\right)\right) - \alpha \nabla_y \phi\left(y^k, x_\alpha\left(y^k\right)\right)$$

where

$$x_\alpha\left(y^k\right) = \arg\max_{x \in \Lambda}\left\{\left\langle F\left(y^k\right), y^k - x\right\rangle - \alpha\phi\left(y^k, x\right)\right\}$$
$$x_\beta\left(y^k\right) = \arg\max_{x \in \Lambda}\left\{\left\langle F\left(y^k\right), y^k - x\right\rangle - \beta\phi\left(y^k, x\right)\right\}$$

Then find

$$d^k = \arg\min\left\{\left[-\nabla \psi_{\alpha\beta}\left(y^k\right)\right]^T y \;\text{ s.t. }\; \|y\| \leq 1\right\}$$

Note that the negative gradient itself may be used as a steepest descent direction so long as it has a bounded norm.

Step 2. Step size determination. Find

$$\theta_k = \arg\min\left\{\psi_{\alpha\beta}\left(y^k + \theta d^k\right) \;\text{ s.t. }\; 0 \leq \theta \leq 1\right\} \tag{5.166}$$

or employ a suitably small constant step size.

Step 2. Stopping test and updating. For $\varepsilon \in \Re^1_{++}$, a preset tolerance, if

$$\left\|\nabla \psi_{\alpha\beta}\left(y^k\right)\right\| < \varepsilon,$$

stop; otherwise set

$$y^{k+1} = y^k - \theta_k d^k\left(y^k\right)$$

and go to Step 1 with k replaced by $k + 1$.

For a numerical example of the gap function method, let us consider $VI(F, \Lambda)$ where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$F(x) = \begin{pmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1 - 5 \\ 0.1x_1x_2 + x_2 - 5 \end{pmatrix}$$

$$\Lambda = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 - 1 \leq 0\}$$

In this example, we employ a D-gap function of the form

$$\psi_{\alpha\beta}(y) = \zeta_\alpha(y) - \zeta_\beta(y)$$

where

$$\zeta_\alpha(y) = \max_{x \in \Lambda} \left\{ \langle F(y), y - x \rangle - \frac{\alpha}{2} \|y - x\|^2 \right\}$$
$$\zeta_\beta(y) = \max_{x \in \Lambda} \left\{ \langle F(y), y - x \rangle - \frac{\beta}{2} \|y - x\|^2 \right\}$$

and $0 < \alpha < \beta$. Thereby, we have defined ϕ to be

$$\phi \equiv \frac{1}{2} \|y - x\|^2$$

Then the gradient information we need is

$$\begin{aligned} \nabla\psi_{\alpha\beta}(y) &= \nabla F(y) \left(x_\beta(y) - x_\alpha(y)\right) + \beta \nabla_y \phi \left(y, x_\beta(y)\right) - \alpha \nabla_y \phi \left(y, x_\alpha(y)\right) \\ &= \begin{pmatrix} 1 & 0 \\ 0.1y_2 & 0.1y_1 + 1 \end{pmatrix} \begin{pmatrix} x_{\beta 1}(y) - x_{\alpha 1}(y) \\ x_{\beta 2}(y) - x_{\alpha 2}(y) \end{pmatrix} + \beta \begin{pmatrix} y_1 - x_{\beta 1}(y) \\ y_2 - x_{\beta 2}(y) \end{pmatrix} \\ &\quad - \alpha \begin{pmatrix} y_1 - x_{\alpha 1}(y) \\ y_2 - x_{\alpha 2}(y) \end{pmatrix} \end{aligned}$$

which leads to

$$\nabla\psi_{\alpha\beta}(y) = \begin{pmatrix} x_{\beta 1}(y) - x_{\alpha 1}(y) + \beta \left(y_1 - x_{\beta 1}(y)\right) - \alpha \left(y_1 - x_{\alpha 1}(y)\right) \\ 0.1y_2 \left(x_{\beta 1}(y) - x_{\alpha 1}(y)\right) + (0.1y_1 + 1) \left(x_{\beta 2}(y) - x_{\alpha 2}(y)\right) + K \end{pmatrix}$$

where

$$K = \beta \left(y_2 - x_{\beta 2}(y)\right) - \alpha \left(y_2 - x_{\alpha 2}(y)\right)$$

Also

$$\begin{aligned} x_\alpha(y) &= \arg\max_{x \in \Lambda} \{\langle F(y), y - x \rangle - \alpha\phi(y, x)\} \\ &= \arg\max_{x \in \Lambda} \left\{ \begin{pmatrix} y_1 - 5 \\ 0.1y_1y_2 + y_2 - 5 \end{pmatrix}^T \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} - \frac{\alpha}{2} \left\| \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} \right\|^2 \right\} \\ &= \arg\max_{x \in \Lambda} \{(y_1 - 5)(y_1 - x_1) + (0.1y_1y_2 + y_2 - 5)(y_2 - x_2) - A\} \end{aligned}$$

$$x_\beta\,(y) = \arg\max_{x\in\Lambda}\{(y_1 - 5)\,(y_1 - x_1) + (0.1y_1y_2 + y_2 - 5)\,(y_2 - x_2) - B\}$$

where

$$A = \frac{1}{2}\alpha\left((x_1 - y_1)^2 + (x_2 - y_2)^2\right)$$
$$B = \frac{1}{2}\beta\left((x_1 - y_1)^2 + (x_2 - y_2)^2\right)$$

If we employ the constant step size $\theta_k = 0.5$, the following table of results is generated:

Iteration k	gap $\psi_{\alpha\beta}(y^k)$	y^k	$x_\alpha\left(y^k\right)$	$x_\beta\left(y^k\right)$
0	0.375	(0, 0)	(0.5000, 0.5000)	(0.5000, 0.5000)
1	2.3512×10^{-2}	(0.3750, 0.3750)	(0.5141, 0.4859)	(0.5035, 0.4965)
2	1.0139×10^{-4}	(0.4750, 0.4635)	(0.5167, 0.4833)	(0.5081, 0.4919)
3	7.005×10^{-6}	(0.5017, 0.4826)	(0.5147, 0.4853)	(0.5108, 0.4892)
4	4.9345×10^{-7}	(0.5095, 0.4866)	(0.5133, 0.4867)	(0.5119, 0.4881)
5	3.5878×10^{-8}	(0.5117, 0.4873)	(0.5123, 0.4872)	(0.5123, 0.4877)
6	2.7672×10^{-9}	(0.5123, 0.4875)	(0.5126, 0.4874)	(0.5124, 0.4876)
7	1.1008×10^{-10}	(0.5124, 0.4875)	(0.5125, 0.4875)	(0.5125, 0.4875)

Evidently, the algorithm terminates with gap $< 10^{-9}$ and approximate solution $y = (0.5125, 0.4875)^T$.

5.7 Other Algorithms for $VI\,(F, \Lambda)$

There are four additional classes of methods for solving finite-dimensional variational inequalities:

1. methods based on differential equations
2. fixed-point methods
3. generalized linear methods
4. succesive linearization and Lemke's algorithm.

Methods based on differential equations express the variational inequality's decision variables as functions of an independent variable t, conveniently called "time",[2] to create differential equations for trajectories that may be continuously deformed to approximate the solution of an equivalent fixed-point problem; the stationary states of these differential equations for $t \longrightarrow \infty$ generate a sequence that converges to the solution of the original variational inequality problem. Fixed-point methods exploit

[2] The independent variable t need not refer to physical time; rather it may be a surrogate for the progress of an algorithm toward the solution of the underlying variational inequality.

the relationship between variational inequalities and fixed-point problems, which enjoy an obvious iterative algoithm

$$x^{k+1} = G(x^k)$$

We have already discussed some aspects of generalized linear methods in Section 5.5. Differential equation and fixed-point methods are discussed by Scarf (1967), Todd (1976), Zangwill and Garcia (1981), and Smith et al. (1997). Generalized linear methods for variational inequalities are reviewed by Pang and Chan (1982), Hammond (1984), and Harker and Pang (1990).

Extensive computational experience during the last decade has produced convincing empirical evidence that a particular method is especially attractive for solving many finite-dimensional variational inequalities. This approach is based on linearization of the nonlinear complementarity formulation of a variational inequality in conjunction with an efficient linear complementarity algorithm – namely Lemke's method. In fact, successive linearization of the nonlinear complementarity formulation of noncooperative equilibria frequently provides the most efficient numerical solution approach. See Cottle et al. (1992) for a discussion of algorithms for linear complementarity problems, as well as Facchinei and Pang (2003a) and Facchinei and Pang (2003b) for additional detail regarding algorithms that exploit the nonlinear complementarity formulation of variational inequalities and Nash equilibria.

5.7.1 *Methods Based on Differential Equations*

Although there are a variety of homotopic differential equations for solution trajectories of variational inequalities, a particularly straightforward approach proposed by Friesz et al. (1994) and Smith et al. (1997) is to equate the rates of change of decision variables to the degree to which the fixed-point equivalent of $VI(F, \Lambda)$ fails to be satisfied, denoted by

$$\Delta \equiv P_\Lambda \{x(t) - \eta F[x(t)]\} - x(t)$$

That is, we write

$$\begin{aligned} \frac{dx(t)}{dt} &= \mu\Delta \\ &= \mu\left[P_\Lambda \{x(t) - \eta F[x(t)]\} - x(t)\right] \qquad (5.167) \\ x(0) &= x^0 \qquad (5.168) \end{aligned}$$

where

$$\mu, \eta \in \Re^1_{++}$$

are parameters adjusted to control stability and assure convergence. It should be apparent that any steady state for which

$$\frac{dx\,(t)}{dt} = 0 \tag{5.169}$$

must correspond to

$$x = P_{\Lambda}\left[x - \eta F\,(x)\right] \tag{5.170}$$

which is recognized, per Lemma 5.2, as the fixed-point equivalent of $VI\,(F, \Lambda)$. Thus, if the dynamics (5.167) and (5.168) lead to (5.169) as $t \longrightarrow \infty$, the desired variational inequality solution is obtained.

5.7.2 *Fixed-Point Methods*

As we have commented before, there is a natural and obvious algorithm associated with any fixed-point problem

$$y = G\,(y)\,,$$

namely

$$y^{k+1} = G\left(y^{k}\right) \tag{5.171}$$

where k is of course the iteration counter. Again we make use of the fact that, for convex feasible regions, $VI\,(F, \Lambda)$ is equivalent to the fixed-point problem $FPP_{\min}\,(F, \Lambda)$; that is,

$$G\,(y) = P_{\Lambda}\left[y - \eta F\,(y)\right] \tag{5.172}$$

where $P_{\Lambda}\,[.]$ is the minimum norm projection operator. It is, therefore, quite reasonable to consider an algorithm for $VI\,(F, \Lambda)$ wherein the iterations follow

$$y^{k+1} = P_{\Lambda}\left[y - \eta F\left(y^{k}\right)\right] \tag{5.173}$$

and

$$\eta \in \Re^{1}_{++}$$

can be considered a step size that may be adjusted to aid convergence. Of course, the righthand side of (5.173) may be expressed as a mathematical program owing to the presence of the minimum norm projection operator. That is, the new iterate y^{k+1} must be the solution of

$$\min_{y \in \Lambda} Z^{k}\,(y) = \frac{1}{2}\left[y^{k} - \eta F\left(y^{k}\right) - y\right]^{T}\left[y^{k} - \eta F\left(y^{k}\right) - y\right] \tag{5.174}$$

$$= \frac{1}{2}\left[\left(y^{k} - y\right) - \eta F\left(y^{k}\right)\right]^{T}\left[\left(y^{k} - y\right) - \eta F\left(y^{k}\right)\right] \tag{5.175}$$

$$= \frac{1}{2}\left\{\left(y^k - y\right)^T \left(y^k - y\right) - 2\eta\left(y^k - y\right)^T F\left(y^k\right) + \eta^2\left[F\left(y^k\right)\right]^T F\left(y^k\right)\right\} \tag{5.176}$$

Upon eliminating the additive constant and multiplying by $(2\eta)^{-1}$, this last expression gives the following form for the subproblems arising in a fixed-point algorithm when the minimum norm projection is involved:

$$\min_{y\in\Lambda}\left(y - y^k\right)^T F\left(y^k\right) + \frac{1}{2\eta}\left(y - y^k\right)^T \left(y - y^k\right) \tag{5.177}$$

which is meant to be solved by an appropriate nonlinear programming algorithm. Browder (1966), Bakusinskii and Poljak (1974), Dafermos (1980), and Bertsekas and Gafni (1982) have used these notions with subtle embellishments to develop algorithms that have linear rates of convergence and perform quite similarly in practice.

5.7.3 *Generalized Linear Methods*

Pang and Chan (1982) offer a very useful and succinct typology of generalized linear methods. In particular, they describe the fundamental subproblem of a generalized linear algorithm to be the following variational inequality

$$F^k(y^{k+1})(x - y^{k+1}) \geq 0 \quad \forall x \in \Lambda$$

which approximates $VI(F, \Lambda)$. Each specific approximation results in a different algorithm. For example, if

$$F^k(y) = F(y^k) + \nabla F(y^k)(y - y^k) \tag{5.178}$$

then the result is Newton's method. If, on the other hand, we use

$$F^k(y) = F(y^k) + \left[\nabla F(y^k)\right]^T (y - y^k) \tag{5.179}$$

then the result is the linearized Jacobi method.

Generalized linear algorithms for variational inequalities are described in some detail by Harker (1988) and Harker and Pang (1990). As we have described above, algorithms belonging to this class proceed by creating a linear approximation of the function $F(x)$ in the variational inequality. The resulting quadratic program can be approached in a variety of ways, including decomposition methods that exploit special structure. Details and applications of generalized linear methods are described by Pang and Chan (1982), Dafermos (1983), Harker (1983), Hammond (1984), Friesz et al. (1985), Nagurney (1987), and Goldsman and Harker (1990).

5.7.4 Successive Linearization with Lemke's Method

We have shown that for certain regularity conditions a variational inequality may be expressed as a nonlinear complementarity problem. Assume that we have a test solution x^k for the equivalent $NCP(F)$ and the function $F(\cdot)$ is continuously differentiable. We approximate F by the first two terms of a Taylor series expansion:

$$F^k(y) = F(y^k) + \left[\text{diag}\nabla F\left(y^k\right)\right]^T (y - y^k)$$

which yields the following linear complimentarity problem, denoted by $LCP\left(F^k\right)$:

$$\begin{aligned}\left[F^k(x)\right]^T x &= 0\\ F^k(x) &\geq 0\\ x &\geq 0\end{aligned}$$

The structure of the successive linearization method is as follows:

Successive Linearization for NCP(F)

Step 0. Initialization. Initialization. Determine an initial feasible solution $x^0 \in \Re^n_+$ and set $k = 0$.

Step 1. Solve the approximating LCP. Approximate the function F about the current solution x^k and, using Lemke's method, solve

$$\begin{aligned}\left[F^k(x)\right]^T x &= 0\\ F^k(x) &\geq 0\\ x &\geq 0\end{aligned}$$

where

$$F^k(x) \equiv F\left(x^k\right) + \nabla F(x^k)\left(x - x^k\right)$$

Call the solution x^{k+1}.

Step 2. Stopping test and updating. For $\varepsilon \in \Re^1_{++}$, a preset tolerance, if

$$\left\|x^{k+1} - x^k\right\| < \varepsilon,$$

stop; otherwise set $k = k + 1$ and go to Step 1.

Convergence of the successive linearization algorithm for nonlinear complementarity problems is treated by Pang and Chan (1982).

5.8 Computing Network User Equilibria

In Section 5.3.6 we discussed the Nash-like equilibrium known as user equilibrium and saw that such problems may be formulated as nonlinear complementarity problems or as variational inequalities. In this section we employ the following numerical example of user equilibrium to illustrate a fixed-point algorithm.

For our example of a user equilibrium let us consider the network of Figure 5.1, consisting of 5 arcs and 4 nodes. For this example, the set of OD pairs is a singleton: $\mathcal{W} = \{(1,4)\}$; as a consequence there are three paths belonging to the set $\mathcal{P} = \mathcal{P}_{14} = \{p_1, p_2, p_3\}$, namely

$$p_1 = \{1,4\} \quad p_2 = \{2,3,4\} \quad p_3 = \{2,5\}$$

In addtion we assume the travel cost for each arc is of the form

$$c_a = A_a + B_a (f_a)^2 \qquad a = 1,2,3,4,5$$

where f_a denotes the total flow on arc a. Moreover, the arc flows and the path flows obey the following relationships:

$$\begin{aligned} f_1 &= h_{p_1} \\ f_2 &= h_{p_2} + h_{p_3} \\ f_3 &= h_{p_2} \\ f_4 &= h_{p_1} + h_{p_2} \\ f_5 &= h_{p_3} \end{aligned}$$

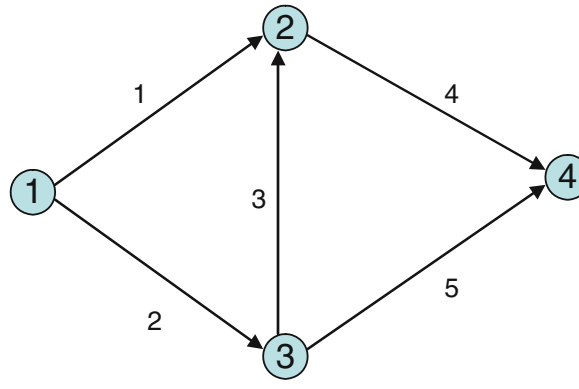


Fig. 5.1 A Simple travel network with five arcs and four nodes

Table 5.1 Parameters

a	A_a	B_a
1	25.0	0.010
2	25.0	0.010
3	75.0	0.001
4	25.0	0.010
5	25.0	0.010

The numerical values for the coefficients A_a and B_a are given in Table 5.1. Furthermore, assuming path costs are additive in arc costs, we may write

$$\begin{aligned} c_{p_1} &= c_1 + c_4 \\ c_{p_2} &= c_2 + c_3 + c_4 \\ c_{p_3} &= c_2 + c_5 \end{aligned}$$

Finally, we stipulate the fixed travel demand

$$T_{14} = 100$$

so that the relevant variational inequality formulation of this particular user equilibrium $UE(C, \Upsilon)$ is: find the traffic pattern $h^* \in \Upsilon$ such that

$$[C\ (h)]^T \left(h - h^*\right) \geq 0 \quad \forall h \in \Upsilon \tag{5.180}$$

where

$$C\ (h) = \begin{bmatrix} A_1 + B_1 \left(h^*_{p_1}\right)^2 + A_4 + B_4 \left(h^*_{p_1} + h^*_{p_2}\right)^2 \\ A_2 + B_2 \left(h^*_{p_2} + h^*_{p_3}\right)^2 + A_3 + B_3 \left(h^*_{p_2}\right)^2 + A_4 + B_4 \left(h^*_{p_1} + h^*_{p_2}\right)^2 \\ A_2 + B_2 \left(h^*_{p_2} + h^*_{p_3}\right)^2 + A_5 + B_5 \left(h^*_{p_3}\right)^2 \end{bmatrix}$$

$$h = \begin{pmatrix} h_{p_1} \\ h_{p_2} \\ h_{p_3} \end{pmatrix}$$

$$\Upsilon = \left\{ \left(h_{p_1}, h_{p_2}, h_{p_3}\right)^T : h_{p_1} + h_{p_2} + h_{p_3} = T_{14} \text{ and } h_{p_1}, h_{p_2}, h_{p_3} \geq 0 \right\}$$

Evidently, feasible region Υ is convex.

For the problem at hand, $UE(C, \Upsilon)$ given by (5.180) takes the following form:

$$C_{p_1}\left(h^*\right)\left(h_{p_1} - h^*_{p_1}\right) + C_{p_2}\left(h^*\right)\left(h_{p_2} - h^*_{p_2}\right) + C_{p_3}\left(h^*\right)\left(h_{p_3} - h^*_{p_3}\right) \geq 0$$

Since $h^*, h \in \Upsilon$, we consider the substitutions

$$\begin{aligned} h^*_{p_1} &= T_{14} - \left(h^*_{p_2} + h^*_{p_3}\right) \\ h_{p_1} &= T_{14} - \left(h_{p_2} + h_{p_3}\right) \end{aligned}$$

Consider the expression

$$\mathbb{Z} \equiv \left[C(h^*)\right]^T (h - h^*) = C_{p_1}(h^*)\left(\left(T_{14} - \left(h_{p_2} + h_{p_3}\right)\right) - \left(T_{14} - \left(h^*_{p_2} + h^*_{p_3}\right)\right)\right) + C_{p_2}(h^*)\left(h_{p_2} - h^*_{p_2}\right) + C_{p_3}(h^*)\left(h_{p_3} - h^*_{p_3}\right) \geq 0$$

We find that

$$\begin{aligned}
\mathbb{Z} &= C_{p_1}(h^*)\left(-\left(h_{p_2} + h_{p_3}\right) + \left(h^*_{p_2} + h^*_{p_3}\right)\right) \\
&\quad + C_{p_2}(h^*)\left(h_{p_2} - h^*_{p_2}\right) + C_{p_3}(h^*)\left(h_{p_3} - h^*_{p_3}\right) \\
&= -C_{p_1}(h^*)\left(h_{p_2} - h^*_{p_2}\right) - C_{p_1}(h^*)\left(h_{p3} - h^*_{p_3}\right) \\
&\quad + C_{p_2}(h^*)\left(h_{p_2} - h^*_{p_2}\right) + C_{p_3}(h^*)\left(h_{p_3} - h^*_{p_3}\right) \\
&= \left(-C_{p_1}(h^*) + C_{p_2}(h^*)\right)\left(h_{p_2} - h^*_{p_2}\right) \\
&\quad + \left(-C_{p_1}(h^*) + C_{p_3}(h^*)\right)\left(h_{p_3} - h^*_{p_3}\right)
\end{aligned}$$

Now the variational inequality may be rewritten as follows: find

$$h^{*T} = \left(h^*_{p_2}\ h^*_{p_3}\right)^T \geq 0$$

such that

$$\begin{bmatrix} -C_{p_1}(h^*) + C_{p_2}(h^*) \\ -C_{p_1}(h^*) + C_{p_3}(h^*) \end{bmatrix}^T \begin{bmatrix} h_{p_2} - h^*_{p_2} \\ h_{p_3} - h^*_{p_3} \end{bmatrix} \geq 0$$

for all

$$h^T = \left(h_{p_2}\ \ h_{p_3}\right)^T \geq 0$$

The corresponding fixed-point iterative scheme is

$$\begin{bmatrix} h^{k+1}_{p_2} \\ h^{k+1}_{p_3} \end{bmatrix} = \begin{bmatrix} h^k_{p_2} - \alpha\left\{-C_{p_1}\left(h^k\right) + C_{p_2}\left(h^k\right)\right\} \\ h^k_{p_3} - \alpha\left\{-C_{p_1}\left(h^k\right) + C_{p_3}\left(h^k\right)\right\} \end{bmatrix}_+$$

for any $\alpha > 0$, where

$$[v]_+ = \max(0, v)$$

Table 5.2 contains a record of iterations corresponding to this fixed-point computational scheme with starting solution

$$h^0_{p_1} = 30,\ \ h^0_{p_2} = 50,\ \ h^0_{p_3} = 20$$

We see that the solution is

$$h^*_{p_1} = 50,\ \ h^*_{p_2} = 0,\ \ h^*_{p_3} = 50$$

Table 5.2 Iterations

Iteration	h_{p_1}	h_{p_2}	h_{p_3}	Error
0	30	50	20	
1	43	16	41	42.0238
2	50.2	0	49.8	19.6286
3	50.04	0	49.96	0.2263
4	50.008	0	49.992	0.0453
5	50.0016	0	49.9984	0.0091
6	50.0003	0	49.9997	0.0018
7	50.0000	0	50.0000	0.0002

The corresponding path costs are

$$c^*_{p_1} = 100,\ c^*_{p_2} = 175,\ c^*_{p_3} = 100\,,$$

indicating a user equilibrium has been obtained.

5.9 Exercises

1. Give a variational inequality statement of user equilibrium with fixed demand that involves only nonnegativity constraints.
2. Prove the existence of a generalized Nash equilibrium $GNE(\Theta, \Omega)$ for suitable regularity conditions.
3. Solve the example of Section 5.5 using a D-gap function.
4. Prove the existence of user equilibrium $UE(C, \Upsilon)$ for fixed, bounded travel demand and continuous cost functions.
5. Prove the uniqueness of user equilibrium $UE(c, \Upsilon_0)$ where c is the vector of arc costs and

$$\Upsilon_0 = \{f : h \geq 0, \Gamma h = T,\ f = \Delta h\}$$
$$\Gamma = \left(\gamma_{ij}^p\right) \text{ is the OD-path incidence matrix}$$
$$\Delta = \left(\delta_{ap}\right) \text{ is the arc-path incidence matirx}$$

   Assume $c(f)$ is strictly monotonically increasing.
6. State a nonlinear program based on Lemma 5.3 for finding user equilibrium arc flows, along with conditions that assure its objective function is single valued.

List of References Cited and Additional Reading

Auchmuty, G. (1989). Variational principles for variational inequalities. *Numerical Functional Analysis and Optimization 10*, 863–874.

Auslender, A. (1976). *Optimisation: Méthodes Numériques*. Paris: Masson.

Bakusinskii, A. B. and B. T. Poljak (1974). On the solution of variational inequalities. *Soviet Mathematics Doklady 15*, 1705–1710.

Bertsekas, D. P. and E. M. Gafni (1982). Projection methods for variational inequalities with application to the traffic assignment problem. *Mathematical Programming Study 17*, 139–159.

Browder, F. E. (1966). Existence and approximation of solutions of nonlinear variational inequalities. *Proceedings of the National Academy of Sciences 56*, 1080–1086.

Cottle, R. W., J. S. Pang, and R. E. Stone (1992). *The Linear Complementarity Problem*. Boston: Academic Press.

Dafermos, S. C. (1980). Traffic equilibrium and variational inequalities. *Transportation Science 14*, 42–54.

Dafermos, S. C. (1983). An iterative scheme for variational inequalities. *Mathematical Programming 26*, 40–47.

Facchinei, F. and J.-S. Pang (2003a). *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volume I. New York: Springer-Verlag.

Facchinei, F. and J.-S. Pang (2003b). *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volume II. New York: Springer-Verlag.

Fiacco, A. V. and G. P. McCormick (1990). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Reprint. Philadelphia: Society for Industrial and Applied Mathematics.

Friesz, T., D., N. Bernstein, R. T. Mehta, and S. Ganjalizadeh (1994). Day-to-day dynamic network disequilibrium and idealized driver information systems. *Operations Research 42*, 1120–1136.

Friesz, T. L., P. A. Viton, and R. L. Tobin (1985). Economic and computational aspects of freight network equilibrium: a synthesis. *Journal of Regional Science 25*, 29–49.

Fukushima, M. (1992). Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming 53*, 99–110.

Goldsman, L. and P. T. Harker (1990). A note on solving general equilibrium problems with variational inequality techniques. *Operations Research Letters 9*, 335–339.

Hammond, J. H. (1984). *Solving asymmetric variational inequality problems and systems of equations with generalized nonlinear programming algorithms*. Ph. D. thesis, Department of Mathematics, MIT.

Harker, P. T. (1983). *Prediction of intercity freight flows: theory and application of a generalized spatial price equilibrium model*. Ph. D. thesis, University of Pennsylvania.

Harker, P. T. (1988). Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational ineqaulities. *Mathematical Programming 41*, 29–59.

Harker, P. T. and J.-S. Pang (1990). Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms, and applications. *Mathematical Programming 48*, 161–220.

Nagurney, A. (1987). Competitive equilibrium problems, variational inequalities, and regional science. *Journal of Regional Science 27*, 55–76.

Ortega, J. M. and W. C. Rheinboldt (2000). *Iterative Solution of Nonlinear Equations in Several Variables* (Reprint ed.). Society for Industrial and Applied Mathematics.

Pang, J.-S. and D. Chan (1982). Iterative methods for variational and complementary problems. *Mathematical Programming 24*, 284–313.

Peng, J.-M. (1997). Equivalence of variational inequality problems to unconstrained minimization. *Mathematical Programming 78*, 347–355.

Scarf, H. E. (1967). The approximation of fixed points of a continuous mapping. *SIAM Journal of Applied Mathematics 15*, 1328–1343.

Smith, T. E., T. L. Friesz, D. Bernstein, and Z. Suo (1997). A comparison of two minimum norm projective dynamic systems and their relationship to variational inequalities. In M. Ferris and J. S. Pang (Eds.), *Complementarity and Variational Problems*, pp. 405–424. SIAM.

Tobin, R. L. (1986). Sensitivity analysis for variational inequalities. *Journal of Optimization Theory and Applications 48*, 191–204.

Todd, M. J. (1976). *The Computation of Fixed Points and Applications*. New York: Springer-Verlag.

Wu, J. H., M. Florian, and P. Marcotte (1993). A general descent framework for the monotone variational inequality problem. *Mathematical Programming 61*, 281–300.

Yamashita, N., K. Taji, and M. Fukushima (1997). Unconstrained optimization reformulations of variational inequality problems. *Journal of Optimization Theory and Applications 92*(3), 439–456.

Zangwill, W. I. and C. B. Garcia (1981). *Pathways to Solutions, Fixed Points, and Equilibria*. Englewood Cliffs, NJ: Prentice-Hall.

Chapter 6
Differential Variational Inequalities and Differential Nash Games

In this chapter we focus on extending the notion of a noncooperative Nash equilibrium to a dynamic, continuous-time setting. The dominant mathematical perspective we will employ is that of a *differential variational inequality*. In fact we shall see that many of the results obtained in the previous chapter for finite-dimensional variational inequalities and static games carry over with some slight modifications to the dynamic, continuous-time setting we now address.

The dynamic games we shall exclusively consider will be deterministic; that is, there is no uncertainty. The solution concept we employ for these games is that of a Nash equilibrium, appropriately generalized from the static setting of Chapter 5 to the dynamic setting of the present chapter. Furthermore, the dynamic games we consider are known as *open-loop* games. An open-loop game is one for which initial information is perfect and complete solution trajectories from the start time t_0 to the end time t_f can be calculated, without reliance on any feedback. By contrast, *closed-loop* games involve the explicit consideration of feedback; we shall not consider closed-loop games in this book.

The following is a preview of the principal topics covered in this chapter:

Section 6.1: Infinite-Dimensional Variational Inequalities. We introduce the infinite-dimensional variational inequality problem and explore the existence of solutions to it.

Section 6.2: Differential Variational Inequalities. We define a deterministic differential variational inequality and establish necessary conditions that must be satisfied by any solution to it.

Section 6.3: Differential Nash Games. We define the notion of a differential Nash game. We note that such games may be expressed as differential variational inequalities.

Section 6.4: Fixed-Point Algorithm. We present and study a simple fixed point algorithm that is often very effective for solving infinite-dimensional and differential variational inequalities.

Section 6.5: Descent in Hilbert Space with Gap Functions. We study the solution of infinite-dimensional and differential variational inequalities using gap functions that are minimized by a descent method based on gradient projection.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_6, © Springer Science+Business Media, LLC 2010

Section 6.6: Differential Variational Inequalities with Time Shifts. In preparation for subsequent applications that are Nash games involving explicit time shifts, we derive necessary conditions for differential variational inequalities involving time shifts.

6.1 Infinite-Dimensional Variational Inequalities

In Chapter 4 we encountered infinite-dimensional variational inequalities as necessary conditions for infinite-dimensional mathematical programs. However, infinite-dimensional variational inequalities may be studied even when they are not necessary conditions of some mathematical program. In particular, if the vector space V is a topological vector space and $F : U \times \Re^1_+ \longrightarrow V$, where $U \subseteq V$, we may pose the problem

$$\left.\begin{array}{c} u^* \in U \\ \langle F(u), u - u^* \rangle \geq 0 \quad \forall u \in U \end{array}\right\} \quad VI(F, U) \tag{6.1}$$

which is notationally similar to the variational inequalities considered in Chapter 5. In fact we may also define fixed-point problems and nonlinear complmentarity problems relative to $F(.)$ and U. These may be solved by adaptation of the numerical methods presented in Chapter 5 to infinite-dimensional spaces. Moreover, nearly every result presented in Chapter 5 regarding the relationship of fixed-point problems, nonlinear complementarity problems, variational inequalities and mathematical programs may be proven for either general or specific infinite-dimensional spaces.

In this chapter, we are interested in a specific class of infinite-dimensional variational inequalities, namely so-called differential variational inequalities. Differential variational inequalities are defined formally in the next section; for now it is enough to recognize two things about them:

1. differential variational inequalities are characterized by explicit state dynamics and explicit controls; and
2. at times we will restate differential variational inequalities as infinite-dimensional variational inequalities without explicit state dynamics, by using the notion of a state operator, to be defined below in Section 6.2.1.

Exploitation of the state operator allows us to easily apply available theory on the existence of solutions to variational inequalities to study the existence of solutions to differential variational inequalities.

Thus, it is appropriate to now present three existence theorems: one for fixed-point problems defined on a simplex in $\Re^n$, one for fixed-point problems in topological vector spaces, and one for variational inequalities in topological vector spaces. The classical result by Brouwer (1910) is presented without proof. The other two theorems and their proofs are due to Browder (1968). The Brouwer theorem in its so-called classical form is:

Theorem 6.1. *Brouwer's classical fixed-point theorem. Under a continuous mapping* $f : S \longrightarrow S \subset \Re^n$ *of an n-dimensional simplex into itself, there exists at least one point* $x \in S$ *such that* $f(x) = x$.

Proof. See Todd (1976). ■

Theorem 6.2. *Browder's elementary fixed-point theorem. Let* U *be a nonempty compact convex subset of a topological vector space* V. *Let* F *be a mapping of* U *into* 2^U. *For each* $u \in U$, $F(u)$ *is a nonempty convex subset of* U. *Suppose further that for each* $v \in U$, $F^{-1}(v) = \{u : u \in U, v = F(u)\}$ *is open in* U. *Then there exists* $\bar{u} \in U$ *such that* $\bar{u} \in F(\bar{u})$.

Proof. We follow Browder (1968). For any $v \in U$, we note that $F^{-1}(v)$ is an open subset of U. Moreover, each point x of U lies in at least one such open subset. Because U is compact, there exists a finite family $\{v_1, v_2, \ldots, v_n\}$ such that each $v_j \in U$ where

$$U = \bigcup_{j=1}^{n} F^{-1}(v_j)$$

Take $\{\beta_1, \beta_2 \ldots, \beta_n\}$ to be a partition of unity corresponding to the above covering. In particular, let each β_j be a continuous mapping of U into $\Re^1$ such that

$$\beta_j(u) = 0 \quad \forall u \notin F^{-1}(v_j) \quad j \in \{1, 2, \ldots, n\}$$

while

$$\sum_{j=1}^{n} \beta_j(u) = 1, \;\; 0 \leq \beta_j(u) \leq 1 \quad \forall u \in U \quad j \in \{1, 2, \ldots, n\}$$

Next define a continuous mapping $p : U \longrightarrow U$ by setting

$$p(u) = \sum_{j=1}^{n} \beta_j(u) v_j$$

Because $v_j \in U$, the convex combination $p(u)$ also lies in U. Furthermore, for each j such that $\beta_j(u) \neq 0$, $u \in F^{-1}(v_j)$ so that $v_j \in F(u)$. Consequently, $p(u)$ is a convex linear combination of points in the convex set $F(u)$ and, therefore, it must be that $p(u) \in F(u)$ for each $u \in U$.

Now take $U_0 \subseteq U$ to be the finite-dimensional simplex spanned by the n points $\{v_1, \ldots, v_n\}$. Note that the topology induced on any finite-dimensional subspace of V by the topological structure of V coincides with the usual Euclidean topology. Hence U_0 is homeomorphic to a Euclidean ball. Since $p(\cdot)$ maps U_0 into U_0, by Brouwer's classical fixed-point theorem, $p(\cdot)$ has a fixed point $\bar{u} \in U_0$. That is

$$\bar{u} = p(\bar{u})$$

Because $p(\bar{u}) \in F(\bar{u})$, we have

$$\bar{u} = p(\bar{u}) \in F(\bar{u})$$

This completes the proof. ■

The result on existence of solutions to infinite-dimensional variational inequalities is:

Theorem 6.3. *Browder's fixed-point theorem for infinite-dimentional variational inequalities. Let U be a compact convex subset of the locally convex topological vector space V and F a continuous (single-valued) mapping of U into V^* (the dual space of V). Then there exists $u^* \in U$ such that*

$$\left\langle F\left(u^*\right), u^* - u\right\rangle \geq 0$$

for all $u \in U$.

Proof. Following Browder (1968), we provide a proof by contradiction. Suppose that, for each $u^* \in U$, there exists an element $u \in U$ such that

$$\left\langle F\left(u^*\right), u^* - u\right\rangle < 0 \tag{6.2}$$

For each $u^* \in U$, let

$$G\left(u^*\right) = \left\{u : u \in U, \left\langle F\left(u^*\right), u^* - u\right\rangle < 0\right\} \tag{6.3}$$

Then $G(u^*)$ is nonempty, by the assumption made in (6.2), for each $u^* \in U$. Furthermore $G(u^*)$ is convex for each u^*. Let us now define the function

$$f(u, v) = \langle F(u), u - v\rangle$$

Since $F : U \longrightarrow V^*$ is a continuous mapping of the compact (and hence bounded) set U, $f(u, v)$ is a continuous function of v on U for each fixed $u \in U$. Thus $G^{-1}(u)$ is open in U for each $u \in U$.

By Theorem 6.2, there exists an element $\bar{u} \in U$ such that $\bar{u} \in G(\bar{u})$. However, for this element $\bar{u}$, we have

$$G(\bar{u}) = \{u : u \in U, \langle F(\bar{u}), \bar{u} - u\rangle < 0\}$$

by virtue of (6.3). Since $\bar{u} \in G(\bar{u})$ we have

$$0 > \langle F(\bar{u}), \bar{u} - \bar{u}\rangle = 0$$

which is a contradiction. Hence, supposition (6.2) is false and the theorem is proven. ■

6.2 Differential Variational Inequalities

To articulate an adequately general differential variational inequality with controls, we must specify the function spaces associated with the key mappings that arise in such a problem formulation. The specific function spaces we employ in our exposition are familiar from previous chapters where they allowed optimal control problems to be analyzed as infinite-dimensional mathematical programs. Those same spaces are again employed since we are extending the notion of an optimal control problem to the more general setting of an infinite-dimensional variational inequality.

6.2.1 Problem Definition

We begin by considering the control vector

$$u \in \left(L^2\left[t_0, t_f\right]\right)^m$$

and associated state operator

$$x(u,t) = \arg\left\{\frac{dy}{dt} = f\left(y,u,t\right),\ y\left(t_0\right) = y_0, \Psi\left[y\left(t_f\right), t_f\right] = 0\right\} \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \tag{6.4}$$

where

$$x_0 \in \Re^n \tag{6.5}$$

$$f\ :\ \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n \tag{6.6}$$

$$\Psi\ :\ \Re^n \times \Re^1_+ \longrightarrow \Re^r \tag{6.7}$$

and $\left(L^2\left[t_0, t_f\right]\right)^m$ is the m-fold product of the space of square-integrable functions $L^2\left[t_0, t_f\right]$ with inner product defined by

$$\langle v, u\rangle = \int_{t_0}^{t_f} v^T u dt \tag{6.8}$$

while $\left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ is the n-fold product of the Sobolev space $\mathcal{H}^1\left[t_0, t_f\right]$. The entity $x(u,t)$ is to be interpreted as an operator that tells us the state vector x for each control vector u and each instant of time $t \in \left[t_0, t_f\right] \subset \Re^1_+$ when there are end point conditions which the state variables must satisfy. Working with this operator is, in effect, a supposition that the two-point boundary-value problem involving the state variables has a unique solution for each control vector considered. That is, terminal states obeying the terminal constraints are reachable from the specified initial states for each admissible control. Note that constraints on u are enforced separately, so in working with $x(u,t)$ we are not presuming existence of a solution of the variational inequality to be articulated below. Moreover, unless other conditions are satisfied $x\left(u,t\right)$ is not a solution of the variational inequality

considered in (6.9); rather it should be thought of as a parametric representation of the state vector in terms of the controls. Note also that we do not actually have to explicitly solve for $x(u,t)$, as is made clear in our subsequent analysis. The notion of a state operator $x\,(u,t)$ is precisely that used in Chapter 4 when analyzing optimal control problems from the point of view of infinite-dimensional mathematical programming; this notation is not original to us but has been employed by others; see, for example, Minoux (1986).

Furthermore, we assume that every control vector is constrained to lie in a set

$$U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m,$$

where U is defined to ensure the terminal conditions imposed on the state variables may be reached from the initial conditions intrinsic to (6.4). Given the operator (6.4), the variational inequality of interest to us takes the following form:

$$\left.\begin{array}{c} \text{find } u^* \in U \text{ such that} \\ \left\langle F\left(x(u^*,t),u^*,t\right), u-u^*\right\rangle \geq 0 \quad \forall u \in U \end{array}\right\} \tag{6.9}$$

where

$$F : \left(\mathcal{H}^1\left[t_0,t_f\right]\right)^n \times \left(L^2\left[t_0,t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0,t_f\right]\right)^m$$

Note that, by virtue of the inner product (6.8), we may state the variational inequality (6.9) as

$$\left\langle F\left(x(u^*,t),u^*,t\right), u-u^*\right\rangle \equiv \int_{t_0}^{t_f} \left[F\left(x(u^*),u^*,t\right)\right]^T \left(u-u^*\right) \geq 0$$

We refer to (6.9) as a *differential variational inequality* (with explicit state equations and controls) and give it the symbolic name *DVI(F, f, Ψ, U, x_0, t_0, t_f)*.

6.2.2 Naming Conventions

At this time it is instructive to consider the history of naming conventions for problems like (6.9). In particular, the name *differential variational inequality* has been used by Aubin and Cellina (1984) to describe a somewhat different problem, namely that of finding $x^* \in X \subseteq V$ such that

$$\left\langle \frac{dx}{dt} - g(x), x - x^* \right\rangle \geq 0 \qquad \text{for all } x \in X$$

where V is typically a Hilbert space while $g \in G : X \longrightarrow \Re^n$ is a set valued map. However, the result of generalizing a game with static equilibria to a dynamic game with explicit differential equations of motion is widely referred to as a *differential game*, in accord with the way the name was originally employed by Isaacs (1965).

Since we shall be formulating differential games as variational inequalities with explicit state dynamics and explicit controls, it is natural to call problem (6.9) a *differential variational inequality* (DVI), as we have done.

There are relatively few published applications in which a problem structure like (6.9) arises. Among these few are the papers by Bernstein et al. (1993), Friesz et al. (1993), and Friesz et al. (1996), who have studied problems with a structure such as (6.9), in the context of dynamic traffic assignment. Perakis (2000) and Kachani and Perakis (2002b) used a formulation somewhat similar to *DVI(F, f, Ψ, U, x_0, t_0, t_f)* to study dynamic transportation equilibrium. Kachani and Perakis (2002a) studied certain inventory and supply chain management problems, again using a formulation similar to *DVI(F, f,Ψ, U, x_0, t_0, t_f)*. Kwon et al. (2009) and Mookherjee and Friesz (2008) used the differential variational inequality formalism to study pricing and revenue management problems. The paucity of prior applications exploiting the differential variational formalism developed in this chapter notwithstanding, virtually any dynamic game-theoretic model that views agents to be competing noncooperatively and moving through space and time in a way that maintains a Nash equilibrium with explicit state dynamics and controls may be expressed as a differential variational inequality of the type discussed herein. Subsequent chapters of this book show a differential variational inequality representation is possible for a variety of manufacturing and service operations management applications.

6.2.3 Regularity Conditions for DVI(F, f, Ψ, U, x_0, t_0, t_f)

To analyze (6.9) we will rely on the following notion of regularity:

Definition 6.1. *Regularity of DVI(F, f, Ψ, U, x_0, t_0, t_f). We call DVI(F, f, Ψ, U, x_0, t_0, t_f) regular if:*

R1. $u \in U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m$

R2. $x \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$

R3. $x(u,t) : \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ *exists and is unique, strongly continuous and G-differentiable for all admissible u;*

R4. $\Psi : \Re^n \times \Re^1_+ \longrightarrow \Re^r$ *is continuously differentiable with respect to x and t;*

R5. $F : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow L^2\left[t_0, t_f\right]$ *is continuous with respect to x and u;*

R6. $f : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n$ *is continuously differentiable with respect to x and u;*

R7. $x_0 \in \Re^n$, $t_0 \in \Re^1_+$, *and* $t_f \in \Re^1_{++}$ *are known and fixed;*

R8. $U \subset \left(L^2\left[t_0, t_f\right]\right)^m$ *is convex; and*

R9. there is a constant dual vector $\upsilon \in \Re^r$ *for the terminal constraints* $\Psi\left[x\left(t_f\right), t_f\right] = 0$.

The motivation for this definition of regularity is to parallel as closely as possible those assumptions used in Chapter 4 to analyze traditional optimal control problems from the point of view of infinite-dimensional mathematical programming.

6.2.4 Necessary Conditions

To develop necessary conditions for solutions of *DVI(F, f, Ψ, U, x_0, t_0, t_f)*, we note that (6.9) may be restated as the following optimal control problem

$$\min \upsilon^T \Psi \left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \left[F\left(x^*, u^*, t\right)\right]^T u dt \tag{6.10}$$

subject to

$$\frac{dx}{dt} = f\left(x, u, t\right) \tag{6.11}$$

$$u \in U \tag{6.12}$$

$$x\left(t_0\right) = x_0 \tag{6.13}$$

where $x^* = x(u^*, t)$ is the optimal state vector and $\upsilon \in \Re^r$ is the vector of dual variables for the terminal constraints $\Psi\left[x\left(t_f\right), t_f\right] = 0$. Care must be taken to correctly understand the meaning of optimal control problem (6.10), (6.11), (6.12), and (6.13). In particular, this optimal control problem is a mathematical abstraction and of no use for computation, since its criterion depends on knowledge of the variational inequality solution u^*. Nonetheless, it is valuable for deriving necessary conditions for *DVI(F, f, Ψ, U, x_0, t_0, t_f)*. In particular, the necessary conditions for *DVI(F, f, Ψ, U, x_0, t_0, t_f)* follow directly from the minimum principle and related necessary conditions for (6.10), (6.11), (6.12), and (6.13).

In what follows we will need the Hamiltonian for (6.10), (6.11), (6.12), and (6.13), namely

$$H\left(x, u, \lambda, t\right) = \left[F\left(x^*, u^*, t\right)\right]^T u + \lambda^T f\left(x, u, t\right) \tag{6.14}$$

where $\lambda(t)$ is the adjoint vector that solves the adjoint equations and satisfies the transversality conditions for the given state variables and controls. Note that, for a given state vector and a given instant in time, the expression (6.14) is convex in u when *DVI(F, f, Ψ, U, x_0, t_0, t_f)* is regular in the sense of Definition 6.1. It is now a relatively easy matter to derive the necessary conditions stated in the following theorem:

Theorem 6.4. *Necessary conditions for DVI(F, f, Ψ, U, x_0, t_0, t_f). Consider the differential variational inequality DVI(F, f, Ψ, U, x_0, t_0, t_f) defined by (6.9) with t_0, $x(t_0)$, and t_f fixed. When regularity in the sense of Definition 6.1 holds, necessary conditions for $u^* \in U$ to be a solution are:*

1. the variational inequality:

$$\left[F\left(x^*, u^*, t\right) + \nabla_u \left(\lambda^*\right)^T f\left(x^*, u^*, t\right)\right]^T \left(u - u^*\right) \geq 0 \quad \forall u \in U; \tag{6.15}$$

2. the state initial-value problem:

$$\frac{dx^*}{dt} = f\left(x^*, u^*, t\right) \tag{6.16}$$

$$x^*\left(t_0\right) = x_0; \tag{6.17}$$

3. the adjoint dynamics:

$$(-1)\frac{d\lambda^*}{dt} = \nabla_x\left(\lambda^*\right)^T f\left(x^*, u^*, t\right); \text{ and} \tag{6.18}$$

4. the transversality conditions:

$$\lambda\left(t_f\right) = v^T \frac{\partial \Psi\left[x^*\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)}$$

Proof. The Pontryagin minimum principle is a necessary condition for optimal control problem (6.10) through (6.13). Hence

$$u^* = \arg\left\{\min_{u \in U} H\left(x^*, u, \lambda^*, t\right)\right\} \tag{6.19}$$

for each $t \in \left[t_0, t_f\right]$, which in turn has the necessary condition

$$\left[\nabla_u H\left(x^*, u^*, \lambda^*, t\right)\right]^T \left(u - u^*\right) \geq 0 \quad u, u^* \in U$$

Note that

$$\nabla_u H\left(x, u, \lambda, t\right) = F\left(x^*, u^*, t\right) + \nabla_u \lambda^T f\left(x, u, t\right)$$

where for given u

$$\begin{aligned}
\lambda(u,t) &= \arg\left\{(-1)\frac{d\lambda}{dt} = \nabla_x H\left(x,u,\lambda,t\right),\ \lambda\left(t_f\right) = v^T\frac{\partial\Psi\left[x\left(t_f\right),t_f\right]}{\partial x\left(t_f\right)}\right\} \\
&= \arg\left\{(-1)\frac{d\lambda}{dt} = \nabla_x\left[F\left(x^*,u^*,t\right)\right]^T u + \nabla_x\lambda^T f\left(x,u,t\right),\right. \\
&\qquad \left.\lambda\left(t_f\right) = v^T\frac{\partial\Psi\left[x\left(t_f\right),t_f\right]}{\partial x\left(t_f\right)}\right\} \\
&= \arg\left\{(-1)\frac{d\lambda}{dt} = \nabla_x\lambda^T f\left(x,u,t\right),\ \lambda\left(t_f\right) = v^T\frac{\partial\Psi\left[x\left(t_f\right),t_f\right]}{\partial x\left(t_f\right)}\right\}
\end{aligned} \tag{6.20}$$

since $x(u,t)$ is completely determined by knowledge of the controls u. The theorem follows immediately. ■

6.2.5 Existence

We are ready to state and prove the following existence result:

Theorem 6.5. *Existence of a solution to DVI(F, f, Ψ, U, x_0, t_0, t_f). When regularity in the sense of Definition 6.1 and U is compact, DVI(F, f, Ψ, U, x_0, t_0, t_f) has a solution.*

Proof. By the assumption of regularity $x(u,t)$ is well defined and continuous. So $F\left(x(u,t),u,t\right)$ is continuous in u. Also, by regularity, we know U is convex and compact. Consequently, by Theorem 6.3, *DVI(F, f, Ψ, U, x_0, t_0, t_f)* has a solution. ■

6.2.6 Nonlinear Complementarity Reformulation

In this section we view the set of feasible controls U as arising from linear constraints; that is

$$U = \{u \geq 0 : Au \leq b\}$$

where $b \in \Re^{\ell}$ is a constant vector and

$$A = \left(a_{ij}\right)$$

is a constant $\ell \times m$ matrix. We restate *DVI(F, f, Ψ, U, x_0, t_0, t_f)* by converting it into a nonlinear complementarity problem in infinite dimensions. This manipulation is accomplished by examining the Kuhn-Tucker conditions for the finite-dimensional variational inequality principle of Theorem 6.4:

$$\nabla_u H\left(x,u,\lambda,t\right) + \sum_{j=1}^{m} \rho_j \nabla_u \left(-u_j\right) + \sum_{j=1}^{\ell} \zeta_j \nabla_u \left(Au-b\right)_j = 0 \tag{6.21}$$

$$\begin{aligned} \rho_j u_j &= 0 \\ \rho_j &\geq 0 \\ \zeta_j \left(Au-b\right)_j &= 0 \\ \zeta_j &\geq 0 \end{aligned} \tag{6.22}$$

for all $j \in [1,m]$. So for each $i = 1,2,\ldots,m$

$$F_i\left(x,u,t\right) + \sum_{j=1}^{n} \lambda_j \frac{\partial}{\partial u_i} f_j\left(x,u,t\right) + \sum_{j=1}^{\ell} \zeta_j \frac{\partial}{\partial u_i} \left(Au-b\right)_j = \rho_i \geq 0 \tag{6.23}$$

or

$$F_i\left(x,u,t\right) + \sum_{j=1}^{n} \lambda_j \frac{\partial}{\partial u_i} f_j\left(x,u,t\right) + \sum_{j=1}^{\ell} a_{ji} \zeta_j = \rho_i \geq 0$$

Since $\rho_i u_i = 0$ because of complementary slackness, we have

$$\left[F_i(x,u,t) + \sum_{j=1}^{n} \lambda_j \frac{\partial}{\partial u_i} f_j(x,u,t) + \sum_{j=1}^{\ell} a_{ji}\zeta_j \right] u_i = \rho_i u_i = 0 \tag{6.24}$$

for each $i = 1, 2, \ldots, m$. Thus, we arrive at the following functional nonlinear complementarity problem:

$$\langle G(z,t), z \rangle = 0 \tag{6.25}$$
$$G(z,t) \geq 0 \tag{6.26}$$
$$z \geq 0 \tag{6.27}$$

where

$$z \in \left(L^2\left[t_0, t_f\right]\right)^{2m+\ell}$$
$$G : \left(L^2\left[t_0, t_f\right]\right)^{2m+\ell} \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^{2m+\ell}$$

and

$$G = \begin{pmatrix} F(x,u,t) + [\nabla_u f(x,u,t)]^T \cdot \lambda + [\nabla_u (Au-b)]^T \cdot \zeta \\ Au - b \\ u \end{pmatrix} \quad z = \begin{pmatrix} u \\ \zeta \\ \rho \end{pmatrix}$$

for which it is understood that x and λ are operators obeying (6.4) and (6.20). Note that in the event $F(.,.,.)$ and $f(.,.,.)$ are linear in u, the reformulation (6.25), (6.26), and (6.27) yields a linear complimentarity problem.

6.3 Differential Nash Games

In this section we want to develop definitions and formulations of dynamic games that employ generalizations of the notions of Nash and generalized Nash equilibria, familiar from Chapter 5, as solution concepts. As previously noted, we will be solely concerned with *open-loop* games. Recall that, an open-loop game is one for which initial information is perfect and complete solution trajectories from the start time t_0 to the end time t_f can be calculated, without reliance on any feedback.

6.3.1 Differential Nash Equilibrium

We need to stipulate that each agent $i \in [1, N]$ has its own control and own state tuples, namely x^i and $u^i \in \Omega_i$, where Ω_i is the set of admissible controls for

agent $i \in [1, N]$. The non-own control and non-own state vectors faced by agent $i \in [1, N]$ are u^{-i} and x^{-i} where

$$u = \begin{pmatrix} u^i \\ u^{-i} \end{pmatrix} \quad x = \begin{pmatrix} x^i \\ x^{-i} \end{pmatrix}$$

for each partition of variables into own and non-own tuples. We will employ the following definition of a differential Nash equilibrium:

Definition 6.2. *Differential Nash equilibrium. Suppose there are N agents, each of which chooses a feasible strategy vector u^i from the strategy set Ω_i which is independent of the other players' strategies. Furthermore, every agent $i \in [1, N]$ has a cost (disutility) functional $J_i(u) : \Omega \longrightarrow \Re^1$ that depends on all agents' strategies where*

$$\Omega = \prod_{i=1}^{N} \Omega_i$$
$$u = \left(u^i : i = 1, \ldots, N\right)$$

Every agent $i \in [1, N]$ seeks to solve the problem

$$\min J_i(u^i, u^{-i}) = K_i\left[x^i\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \Theta_i(x^i, u^i, x^{-i}, u^{-i}, t)dt \tag{6.28}$$

subject to

$$\frac{dx^i}{dt} = f^i\left(x^i, u^i, t\right) \tag{6.29}$$
$$x^i\left(t_0\right) = x_0^i \tag{6.30}$$
$$\Psi^i\left[x^i\left(t_f\right), t_f\right] = 0 \tag{6.31}$$
$$u^i \in \Omega_i\,, \tag{6.32}$$

for each fixed yet arbitrary non-own control tuple

$$u^{-i} = \left(u^j : j \neq i\right)$$

where x_0^i is a vector of initial values of x^i, the state tuple of the i^{th} agent, and

$$x^{-i} = \left(x^j : j \neq i\right)$$

is the corresponding non-own state tuple. A differential Nash equilibrium is a tuple of strategies u such that each u^i solves the optimal control problem (6.28), (6.29), (6.30), (6.31), and (6.32); that equilibrium is denoted as DNE(Θ, f, K, Ψ, Ω, x_0, t_0, t_f).

In other words, we have the familiar situation wherein no agent may lower his/her cost (disutility) by unilaterally altering his/her strategy. The vectors and mappings intrinsic to Definition 6.2 are the following:

$$\begin{aligned}
u^i &\in \left(L^2\left[t_0, t_f\right]\right)^{m_i} \\
x^i &\in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{n_i} \\
x_0^i &\in \Re^{n_i} \\
m &= m_1 + m_2 + \ldots + m_N \\
n &= n_1 + n_2 + \ldots + n_N \\
\Theta_i &: \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{n} \times \left(L^2\left[t_0, t_f\right]\right)^{m} \times \Re_+^1 \longrightarrow L^2\left[t_0, t_f\right] \\
f^i &: \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{n_i} \times \left(L^2\left[t_0, t_f\right]\right)^{m_i} \times \Re_+^1 \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^{n_i} \\
K^i &: \Re^{n_i} \times \Re_+^1 \longrightarrow \Re^1 \\
\Psi^i &: \Re^{n_i} \times \Re_+^1 \longrightarrow \Re^{r_i}
\end{aligned}$$

for each agent $i \in [1, N]$.

It is intuitive that a differential Nash equilibrium may be represented as a differential variational inequality. In fact, the following result holds:

Theorem 6.6. *Differential variational inequality equivalent to differential Nash equilibrium. Take t_0, $x(t_0)$, and t_f to be fixed. There is a differential variational inequality equivalent to the differential Nash equilibrium DNE(Θ, f, K, Ψ, U, x_0, t_0, t_f) when $f^i\left(x^i, u^i, t\right)$ and $\Theta_i(x^i, u^i, x^{-i}, u^{-i}, t)$ are convex and continuously differentiable with respect to $\left(x^i, u^i\right)$ for all fixed non-own tuples $\left(x^{-i}, u^{-i}\right)$, for all $i \in [1, N]$.*

Proof. The relevant Hamiltonian for each agent $i \in [1, N]$ is

$$H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i}) = \Theta^i(x^i, u^i, x^{-i}, u^{-i}, t) + \left(\lambda^i\right)^T f^i\left(x^i, u^i, t\right)$$

and the minimum principle has, by virtue of convexity, the necessary and sufficient condition

$$\left[\nabla_{u^i} H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i})\right]^T \left(v^i - u^i\right) \geq 0 \text{ for all } v^i \in \Omega_i \tag{6.33}$$

where the λ^i are tuples of adjoint variables determined by

$$\begin{aligned}
(-1)\frac{d\lambda^i}{dt} &= \nabla_{x^i}\left(\lambda^i\right)^T f^i(x, u, t) \\
\lambda^i\left(t_f\right) &= \frac{\partial \Gamma^i\left[x\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)} \\
\Gamma^i\left[x\left(t_f\right), t_f\right] &= K^i\left[x\left(t_f\right), t_f\right] + \left(\gamma^i\right)^T \Psi^i\left[x\left(t_f\right), t_f\right]
\end{aligned}$$

for agent $i \in [1, N]$. Let us define the tuples

$$y^i = \begin{pmatrix} x^i \\ \lambda^i \end{pmatrix} \text{ and } y^{-i} = \begin{pmatrix} x^{-i} \\ \lambda^{-i} \end{pmatrix}$$

$$g^i = \begin{pmatrix} f^i \\ (-1)\nabla_{x^i} \left(\lambda^i\right)^T f^i \left(x, u, t\right) \end{pmatrix}$$

$$\Phi^i \left[x\left(t_f\right), t_f\right] = \begin{pmatrix} \Gamma^i \left[x\left(t_f\right), t_f\right] \\ \lambda^i \left(t_f\right) - \dfrac{\partial \Gamma^i \left[x\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)} \end{pmatrix} = 0$$

for each $i \in [1, N]$, so that

$$y = \left(y^i : i = 1, \ldots, N\right)$$
$$g = \left(g^i : i = 1, \ldots, N\right)$$
$$\Phi = \left(\Phi^i : i = 1, \ldots, N\right)$$

Also

$$y\left(t_0\right) = y_0 = \begin{pmatrix} x\left(t_0\right) \\ \lambda\left(t_0\right) \text{ free} \end{pmatrix}$$

In addition we define

$$G^i(y^i, u^i, t; y^{-i}, u^{-i}) = \nabla_{u^i} H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i}) \tag{6.34}$$

$$G = \left(G^i : i = 1, \ldots, N\right) \tag{6.35}$$

It follows from (6.33) and the above notation that

$$\left.\begin{array}{c} u^* \in \Omega \equiv \prod_{i=1}^N \Omega_i \\ \displaystyle\int_{t_0}^{t_f} \left[G(y\left(u^*, t\right), u^*, t)\right]^T \left(v - u^*\right) dt \geq 0 \quad \forall v \in \Omega \end{array}\right\} \tag{6.36}$$

where

$$y(u, t) = \arg\left\{\frac{dy}{dt} = g\left(y, u, t\right),\ y\left(t_0\right) = y_0,\ \Phi\left[y\left(t_f\right), t_f\right] = 0\right\} \tag{6.37}$$

If given differential variational inequality (6.46) and (6.47), by selecting $v^j = u^{*j}$ for all $j \neq i$, the minimum principle is recovered for each individual $i \in [1, N]$. ■

6.3.2 Generalized Differential Nash Equilibrium

When the strategy set and dynamics of any agent $i \in [1, N]$ depend on non-own strategies u^{-i} and non-own states x^{-i}, extension of the definition of a differential Nash equilibrium to a generalized differential Nash equilibrium is exactly what we would expect. That is, we have the following definition:

Definition 6.3. *Generalized differential Nash equilibrium. Suppose there are N agents, each of which chooses a feasible strategy vector u^i from the strategy set $\Omega_i(u)$ that depends on all agents' strategies where*

$$u = \left(u^i : i = 1, \ldots, N\right)$$

Furthermore, every agent $i \in [1, N]$ has a cost (disutility) functional $J_i(u) : \Omega(u) \longrightarrow \Re^1$ that depends on all agents' strategies where

$$\Omega(u) = \prod_{i=1}^{N} \Omega_i(u)$$

Every agent $i \in [1, N]$ seeks to solve the problem

$$\min J_i(u^i, u^{-i}) = K_i\left[x^i\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \Theta_i(x, u, t)dt \tag{6.38}$$

subject to

$$\frac{dx^i}{dt} = f^i(x, u, t) \tag{6.39}$$

$$x^i(t_0) = x_0^i \tag{6.40}$$

$$\Psi^i\left[x\left(t_f\right), t_f\right] = 0 \tag{6.41}$$

$$u^i \in \Omega_i(u), \tag{6.42}$$

for each fixed yet arbitrary non-own control tuple

$$u^{-i} = \left(u^j : j \neq i\right)$$

where x_0^i is a vector of initial values of x^i, the state tuple of the i^{th} agent, and

$$x^{-i} = \left(x^j : j \neq i\right)$$

is the corresponding non-own state tuple. A generalized differential Nash equilibrium is a tuple of strategies u such that each u^i solves the optimal control problem (6.38), (6.39), (6.40), (6.41), and (6.42) and is denoted as GDNE(Θ, f, K, Ψ, U, x_0, t_0, t_f).

It is straightforward to define the notion of a differential quasivariational inequality; in turn it is possible to show that *GDNE(Θ, f, K, Ψ, U, x₀, t₀, t_f)* is equivalent to a differential quasivariational inequality. Of course, a generalized differential Nash equilibrium may be represented as a differential quasivariational inequality, as the reader may easily verify."

6.4 Fixed-Point Algorithm

In order to apply the results developed above regarding the relationship of dynamic Nash games to differential variational inequalities, we must be able to compute the solutions to differential variational inequalities. It should come as no surprise that there is often an equivalent functional fixed-point problem corresponding to a given differential Nash game. This formulation provides an immediate, simple and sometimes quite effective algorithm for solving *DVI(F, f, Ψ, U, x₀, t₀, t_f)*.

6.4.1 Formulation

In particular, we are now ready to state and prove the following result:

Theorem 6.7. *Fixed-point formulation of DVI(F, f, Ψ, U, x₀, t₀, t_f). When regularity in the sense of Definition 6.1 holds and $f(x,u)$ is convex in (x,u), DVI(F, f, Ψ, U, x₀, t₀, t_f) is equivalent to the following fixed-point problem:*

$$u = P_U \left[u - \alpha F\left(x(u,t), u, t\right)\right]$$

where $P_U[.]$ is the minimum norm projection onto $U \subseteq \left(L^2[t_0,\tau]\right)^m$ and $\alpha \in \Re^1_{++}$ is an arbitrary positive constant.

Proof. The fixed-point problem under consideration requires that

$$u = \arg\min_v \left\{ \frac{1}{2} \left\| u - \alpha F\left(x(u,t), u, t\right) - v \right\|^2 : v \in U \right\} \tag{6.43}$$

where $\alpha \in \Re^1_{++}$ is any positive real scalar. That is, we seek the solution of the optimal control problem

$$\min_v \gamma^T \Psi\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \frac{1}{2}\left[u - \alpha F\left(x,u,t\right) - v\right]^2 dt$$

subject to

$$\frac{dx}{dt} = f\left(x,v,t\right);\ x\left(t_0\right) = x_0$$
$$v \in U$$

where u is treated as fixed for the purpose of projection. The necessary conditions of the above optimal control problem, which are also sufficient by virtue of convexity, require

$$\left[\nabla_v H_1\left(x^*, v^*, \eta^*, t\right)\right]^T (v - v^*) \geq 0 \qquad \forall v \in U \tag{6.44}$$

where

$$H_1(x, v, \eta, t) = \frac{1}{2}[u - \alpha F(x, u, t) - v]^2 + \eta^T f(x, v, t)$$

and for given x and v

$$\eta = \arg\left\{(-1)\frac{d\eta}{dt} = \nabla_x H_1(x, v, \eta, t),\ \eta\left(t_f\right) = \gamma^T \frac{\partial \Psi\left[x\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)}\right\}$$

Note that

$$\nabla_v H_1(x, v, \eta, t) = -u + \alpha F(x, u, t) + v + \nabla_v \eta^T f(x, v, t)$$

Because $u = v$ by virtue of (6.43) we have

$$\nabla_u H_1(x, v, \eta, t) = \alpha F(x, u, t) + \nabla_u \left[\eta^T f(x, u, t)\right] \tag{6.45}$$

Now if we set $\lambda = \dfrac{\eta}{\alpha}$; we have

$$\left[F\left(x^*, u^*, t\right) + \nabla_u \left(\lambda^*\right)^T f\left(x^*, u^*, t\right)\right]^T (u - u^*) \geq 0 \qquad \forall v \in U \tag{6.46}$$

which is identical to the finite-dimensional variational inequality principle of Theorem 6.4. The other optimality conditions are also identical. This completes the proof. ■

6.4.2 The Unembellished Algorithm

Naturally there is an associated fixed-point algorithm based on the iterative scheme

$$u^{k+1} = P_U\left[u^k - \alpha F\left(x\left(u^k, t\right), u^k, t\right)\right]$$

The positive scalar may be chosen empirically to assist convergence and may even be changed as the algorithm progresses. The detailed structure of the fixed-point algorithm is given below:

Fixed-Point Algorithm

Step 0. Initialization. Identify an initial feasible solution $u^0 \in U$ and set $k = 0$.

Step 1. Solve the optimal control subproblem. Solve the following optimal control subproblem:

$$\min_v J^k(v) = \gamma^T \Psi\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \frac{1}{2}\left[u^k - \alpha F\left(x^k, u^k, t\right) - v\right]^2 dt \quad (6.47)$$

subject to

$$\frac{dx}{dt} = f(x, v, t) \quad (6.48)$$

$$x(t_0) = x_0 \quad (6.49)$$

$$v \in U \quad (6.50)$$

Call the solution u^{k+1}.

Step 2. Stopping test. If $\left\|u^{k+1} - u^k\right\| \leq \varepsilon_1$ where $\varepsilon_1 \in \Re^1_{++}$ is a preset tolerance, stop and declare $u^* \approx u^{k+1}$. Otherwise set $k = k + 1$ and go to Step 1.

The convergence of this algorithm is guaranteed under certain conditions by the following result:

Theorem 6.8. *Convergence of the unembellished fixed-point algorithm. When DVI(F, f, Ψ, U, x_0, t_0, t_f) is regular in the sense of Definition 6.1, while additionally $F[x(u), u, t]$ is strongly monotonically increasing and satisfies the Lipschitz condition*

$$\|F(x(u,t), u, t) - F(x(v,t), v, t)\| \leq \kappa_0 \|u - v\|$$

for some $\kappa_0 \in \Re^1_{++}$ and all $u, v \in U$, the fixed-point algorithm presented above converges for appropriate $\alpha \in (0, \bar{\alpha})$.

Proof. The projection operator is nonexpansive; that is, we know

$$\begin{aligned} &\left\|P_U\left[u^k - \alpha F\left(x\left(u^k, t\right), u^k, t\right)\right] - P_U\left[u^* - \alpha F\left(x\left(u^*, t\right), u^*, t\right)\right]\right\| \\ &\leq \left\|\left[u^k - \alpha F\left(x\left(u^k, t\right), u^k, t\right)\right] - \left[u^* - \alpha F\left(x\left(u^*, t\right), u^*, t\right)\right]\right\| \\ &= \left\|\left(u^k - u^*\right) - \alpha\left(F^k - F^*\right)\right\| \end{aligned} \quad (6.51)$$

where

$$F^k \equiv F\left(x\left(u^k, t\right), u^k, t\right)$$

$$F^* \equiv F\left(x\left(u^*, t\right), u^*, t\right)$$

and $u^* \in U$ solves *DVI(F, f, Ψ, U, x*$_0$*, t*$_0$*, t*$_f$*)*. Thus, we may write

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left\|\left(u^k - u^*\right) - \alpha\left(F^k - F^*\right)\right\|^2 \tag{6.52}$$
$$= \left(u^k - u^*\right)^2 - 2\alpha\left(F^k - F^*\right)^T\left(u^k - u^*\right) + \alpha^2\left(F^k - F^*\right)^2$$

The given of strong monoticity requires

$$\langle F^k - F^*, u^k - u^*\rangle \geq \beta\left\|u^k - u^*\right\|^2 \tag{6.53}$$

for some $\beta \in \Re^1_{++}$. It follows from (6.52) and (6.53) that

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left\|u^k - u^*\right\|^2 - 2\alpha\beta\left\|u^k - u^*\right\|^2 + \alpha^2\left\|F^k - F^*\right\|^2 \tag{6.54}$$

Because F is Lipschitz continuous, we know

$$\left\|F^k - F^*\right\|^2 \leq \kappa_0\left\|u^k - u^*\right\|^2 \tag{6.55}$$

for some $\kappa_0 \in \Re^1_{++}$. From (6.54) and (6.55) we have

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left(1 - 2\alpha\beta + \kappa_0\alpha^2\right)\left\|u^k - u^*\right\|^2$$

If we stipulate

$$\zeta = 1 - 2\alpha\beta + \kappa_0\alpha^2 < 1 \tag{6.56}$$

then $\sqrt{\zeta} < 1$ and

$$\left\|u^{k+1} - u^*\right\| \leq \sqrt{\zeta}\cdot\left\|u^k - u^*\right\| < \left\|u^k - u^*\right\| \tag{6.57}$$

which, by the contraction mapping theorem, assures convergence. The upper bound on α consistent with convergence satisfies

$$1 - 2\alpha\beta + \kappa_0\alpha^2 = 1$$

or

$$\alpha\left(-2\beta + \kappa_0\alpha\right) = 0 \Longrightarrow \bar{\alpha} = \frac{2\beta}{\kappa_0} > 0$$

The desired result has been proven. ■

6.4.3 Solving the SubProblems

It is important to realize that the fixed-point algorithm of Section 6.4 can be carried out in continuous time provided we employ a continuous-time representation of the solution of each subproblem (6.47), (6.48), (6.49), and (6.50) from Step 1. This may be done using a continuous-time gradient projection method. For our present circumstances, that algorithm may be stated as

Descent in Hilbert Space Algorithm for Projection Subproblems

Step 0. Initialization. Pick $v^{k,0}(t) \in U$ and set $j = 0$.

Step 1. Finding state variables. Solve the state dynamics

$$\frac{dx}{dt} = f\left(x, v^{k,j}, t\right) \tag{6.58}$$

$$x(t_0) = x_0 \tag{6.59}$$

Call the solution $x^{k,j}(t)$. In the event a discrete-time method is used to solve the state dynamics (6.58) and (6.59), curve fitting is used to obtain the continuous-time state vector $x^{k,j}(t)$.

Step 2. Finding adjoint variables. Solve the adjoint dynamics

$$(-1)\frac{d\lambda}{dt} = \nabla_x H_1^k \Big|_{\substack{v=v^{k,j} \\ x=x^{k,j}}} \tag{6.60}$$

$$\lambda\left(t_f\right) = \gamma^T \frac{\partial \Psi\left[x^{k,j}\left(t_f\right), t_f\right]}{\partial x\left(t_f\right)} \tag{6.61}$$

where

$$H_1^k = \frac{1}{2}\left[u^k - \alpha F\left(x^k, u^k, t\right) - v\right]^2 + \lambda^T f\left(x, v^{k,j}, t\right)$$

Call the solution $\lambda^{k,j}(t)$. In the event a discrete-time method is used to solve the adjoint dynamics (6.60) and (6.61), curve fitting is used to obtain the continuous-time adjoint vector $\lambda^{k,j}(t)$.

Step 3. Finding the gradient. Determine

$$\nabla_v J^{k,j}(t) = \nabla_v H_1^k$$

Step 4. Step determination. For a fixed and suitably small fixed step size

$$\theta_k \in \Re_{++}^1$$

determine

$$v^{k,j+1}(t) = P_U\left[v^{k,j}(t) - \theta_k \nabla_v J^{k,j}\right] \tag{6.62}$$

In the event a discrete-time method is used to solve the above projection subproblem, curve fitting is used to obtain the continuous-time control vector (6.62).

Step 5. Stopping test. For $\varepsilon_2 \in \Re^1_{++}$, a preset tolerance, stop if

$$\left\| v^{k,j+1} - v^{k,j} \right\| < \varepsilon_2$$

and declare $v^{k*} \approx v^{k,j+1}$. Otherwise set $j = j + 1$ and go to Step 1.

The reader is reminded that convergence of the gradient projection algorithm described above for fixed-point subproblems was analyzed in Chapter 4.

6.4.4 Numerical Example

Consider *DVI(F, f, Ψ, U, x_0, t_0, t_f)* with the following specific data:

$$\begin{aligned} u &\in \left(L^2[0,1]\right)^3 \\ x &\in \left(\mathcal{H}^1[0,1]\right)^2 \\ x(0) &= \begin{pmatrix} 1 \\ 0.7 \end{pmatrix} \\ t_0 &= 0 \\ t_f &= 5 \end{aligned}$$

Additionally, the key functions are

$$F(x,u,t) = \begin{pmatrix} F_1(x,u,t) \\ F_2(x,u,t) \\ F_3(x,u,t) \end{pmatrix} = \begin{pmatrix} (x_1)^2 - u_1 + u_2 \\ x_2 - (u_2)^2 - u_3 \\ \frac{1}{10}(x_2)^2 - (u_3)^2 \end{pmatrix}$$

$$f(x,u,t) = \begin{pmatrix} f_1(x,u,t) \\ f_2(x,u,t) \end{pmatrix} = \begin{pmatrix} \frac{1}{5}x_1 + \frac{1}{2}u_1 + \frac{3}{10}u_2 \\ \frac{1}{4}x_2 + \frac{1}{2}u_2 - \frac{1}{5}u_3 \end{pmatrix}$$

and the set of admissible controls is

$$U = \{u : 1 \geq u_1 \geq 0.2,\ 1.2 \geq u_2 \geq 0.2,\ 1.3 \geq u_3 \geq 0.2\}$$

The fixed-point parameter is $\alpha = 0.05$. A fifth-power polynomial was used to express the controls, adjoint variables, and state variables as continuous functions of time. Also the nominal decision time interval is $[0, 5]$. The stopping tolerances

for both fixed-point and descent iterations were set at $\varepsilon = 10^{-2}$. The combined fixed-point-descent algorithm converged after 12 fixed-point iterations; each of the descent subproblems converged in nine or fewer iterations. We forgo the detailed symbolic statement of this example and, instead, provide numerical results in graphical form. Figure 6.1 shows the plot of controls u^* (left) and states x^* (right) against time.

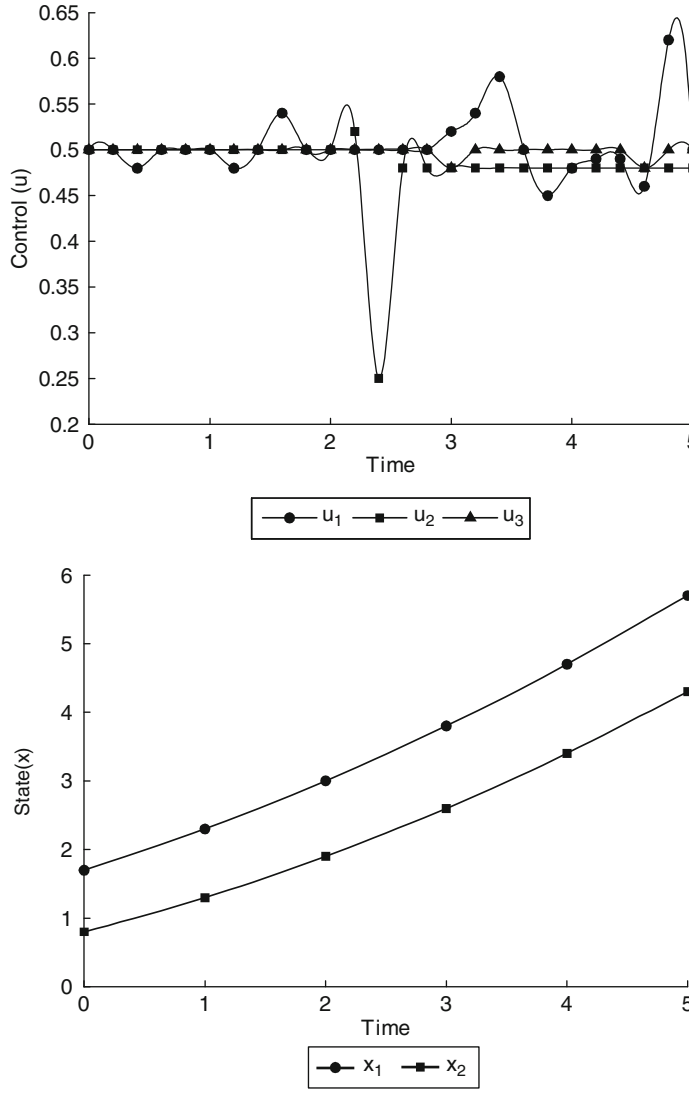


Fig. 6.1 u^* vs. time t and x^* vs. time t plots

6.5 Descent in Hilbert Space with Gap Functions

The unembellished fixed-point algorithm presented in Section 6.4.2 above is neither sophisticated nor completely reliable. In particular, it is known to sometimes converge slowly even when the regularity conditions invoked to assure convergence are satisfied. Of course, it may also fail to converge when those conditions are violated, although not always since the conditions are merely sufficient for convergence. A variety of other algorithms may be employed. One such re-expresses the DVI of interest as a nonlinear complementarity problem in function space. That nonlinear complementaity problem may be approximated through time discretization, then linearized and Lemke's algorithm employed. Another possible approach is to employ a gap function to create an equivalent optimal control problem that may be solved using the tools we developed in Chapter 4 for infinite-dimensional mathematical programs.

Using the notion of a gap function, a variational inequality problem can be converted to an equivalent optimization problem, whose objective function is always nonnegative and whose optimal objective function value is zero if and only if the optimal solution solves the original variational inequality problem. Several algorithms in this class have been developed for finite-dimensional variational inequalities; see, for example, Zhu and Marcotte (1994), Yamashita et al. (1997), Patriksson (1997), and Peng (1997). For infinite-dimensional problems, Zhu and Marcotte (1998) and Konnov et al. (2002) present descent methods using gap functions in Banach spaces and Hilbert spaces, respectively. Moreover, Konnov and Kum (2001) have provided a gap function method for mixed variational inequalities in Hilbert spaces. The discussion of gap functions given next is similar in many respects to the discussion of gap functions for finite-dimensional problems familiar from Chapter 4. However, there are some subtle yet important differences. Furthermore, inclusion of a complete discussion of gap functions here, although somewhat repetitive, helps to make this chapter self-contained.

6.5.1 Gap Functions in Hilbert Spaces

When the regularity conditions given in Definition 6.1 hold, *DVI(F, f, Ψ, U, x_0, t_0, t_f)* belongs to the class of infinite-dimensional variational inequalities considered by Konnov et al. (2002), wherein U is a nonempty closed and convex subset and F is a continuously differentiable mapping of u. This allows us to analyze *DVI(F, f, Ψ, U, x_0, t_0, t_f)* by considering gap functions, which we define as follows:

Definition 6.4. *Gap function defined. A function $G : U \longrightarrow \Re_+$ is called a gap function for DVI(F, f, Ψ, U, x_0, t_0, t_f) when the following statements hold:*

1. $G(u) \geq 0$ *for all* $u \in U$
2. $G(u) = 0$ *if and only if u is the solution of DVI(F, f, Ψ, U, x_0, t_0, t_f).*

In particular, we will consider gap functions of the form

$$G_\alpha(u) = \max_{v \in U} \Phi_\alpha(u, v) \tag{6.63}$$

where

$$\Phi_\alpha(u, v) = \langle F[x(u,t), u, t], u - v\rangle - \alpha\phi(u, v) \tag{6.64}$$

$$x(u,t) = \arg\left\{\frac{dy}{dt} = f(y, u, t),\ y(t_0) = y_0,\ \Psi\left[y\left(t_f\right), t_f\right] = 0\right\}$$

$$\in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \tag{6.65}$$

$$U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m \tag{6.66}$$

$$\alpha \in \Re^1_{++} \tag{6.67}$$

and the function ϕ appearing in (6.63) satisfies the following assumptions:

A1. ϕ is continuously differentiable on $\left(L^2\left[t_0, t_f\right]\right)^{2m}$;
A2. ϕ is nonnegative on $\left(L^2\left[t_0, t_f\right]\right)^{2m}$;
A3. $\phi(u, v) = 0$ if and only if $u = v$; and
A4. $\phi(u, v)$ is strongly convex in $v \in U$ with modulus $c > 0$ for any $u \in \left(L^2\left[t_0, t_f\right]\right)^m$; that is

$$\phi(u, v) + \langle \nabla_v \phi(u, v), u - v\rangle + \frac{1}{2} c \|u - v\|^2 \leq \phi(u, u) = 0$$

for all $u \in U$.

Yamashita et al. (1997) propose, for finite-dimensional spaces, the following ϕ-functions that satisfy the four assumptions listed above:

1. $\phi_1(u, v) = \tau_1(u - v)$, where τ_1 is nonnegative, continuously differentiable, strongly convex, and $\tau_1(0) = 0$;
2. $\phi_2(u, v) = \tau_2(v) - \tau_2(u) - \langle \nabla \tau_2(u), u - v\rangle$, where τ_2 is twice continuously differentiable, and strongly convex; and
3. $\phi_3(u, v) = \langle u - v, M(u)(u - v)\rangle$, where $M(u)$ is a continuously differentiable, symmetric, and uniformly positive-definite matrix.

Konnov and Kum (2001) and Konnov et al. (2002) observe that functions of the type ϕ_2 and ϕ_3, as defined above, may be used to develop gap functions appropriate for Hilbert spaces.

We note, following Konnov and Kum (2001) and Konnov et al. (2002), that the maximization problem (6.63) has a unique solution, since $\Phi_\alpha(u, v)$ is strongly convex in v and U is convex. We will use $v_\alpha(u)$ to denote the solution of (6.63); that is,

$$G_\alpha(u) = \Phi_\alpha[u, v_\alpha(u)]$$

With these preliminaries, we are now able to state and prove the following result:

Lemma 6.1. *Gap function for DVI(F, f, U, Ψ, x_0, t_0, t_f). The function $G_\alpha(u)$ defined by (6.63) is a gap function for DVI(F, f, U, Ψ, x_0, t_0, t_f). In particular, u is the solution to DVI(F, f, U, Ψ, x_0, t_0, t_f), if and only if $u = v_\alpha(u)$.*

Proof. The proof is in two parts.

(i) $\left[u = v_\alpha(u) \Longrightarrow DVI(F, f, \blacksquare, U, x_0, t_0, t_f)\right]$ The optimality condition for (6.63) is

$$\left\langle \frac{\partial \Phi_\alpha(u, v_\alpha)}{\partial v}, v - v_\alpha \right\rangle \leq 0 \qquad \forall v \in U$$

That is

$$\langle -F(x, u, t) - \alpha \nabla_v \phi(u, v_\alpha), v - v_\alpha \rangle \leq 0 \qquad \forall v \in U \tag{6.68}$$

Substituting u for v in (6.68), we obtain

$$\langle F(x, u, t), u - v_\alpha \rangle \geq -\alpha \langle \nabla_v \phi(u, v_\alpha), u - v_\alpha \rangle \qquad \forall u \in U \tag{6.69}$$

Note that strong convexity for the ϕ-function intrinsic to the gap function means the following extension of the tangent line property holds:

$$\phi(u, v_\alpha) + \langle \nabla_v \phi(u, v_\alpha), u - v_\alpha \rangle + \frac{1}{2} c \|u - v_\alpha\| \leq \phi(u, u) \tag{6.70}$$

By virtue of relationships (6.69) and (6.70), we have

$$\begin{aligned} G_\alpha(u) &= \Phi_\alpha(u, v_\alpha) \\ &= \langle F(x, u, t), u - v_\alpha \rangle - \alpha \phi(u, v_\alpha) \\ &\geq -\alpha \langle \nabla_v \phi(u, v_\alpha), u - v_\alpha \rangle - \alpha \phi(u, v_\alpha) \\ &= \alpha [\phi(u, u) - \phi(u, v_\alpha) - \langle \nabla_v \phi(u, v_\alpha), u - v_\alpha \rangle] \\ &\geq \frac{\alpha c}{2} \|v_\alpha - u\|^2 \geq 0 \end{aligned} \tag{6.71}$$

where the property $\phi(u, u) = 0$ is used. Therefore, $G_\alpha(u) \geq 0$ for all $u \in U$. Moreover, if $G_\alpha(u) = 0$, then by (6.71) we have $u = v_\alpha$. From (6.69), we see by inspection that $u = v_\alpha$ solves $DVI(F, f, U, \Psi, x_0, t_0, t_f)$.

(ii) $\left[DVI(F, f, U, \Psi, x_0, t_0, t_f) \Longrightarrow u = v_\alpha(u)\right]$ Suppose now that u is a solution of $DVI(F, f, U, \Psi, x_0, t_0, t_f)$. Then

$$\langle F(x, u, t), v - u \rangle \geq 0 \qquad \forall v \in U$$

and it follows that

$$\Phi_\alpha(u, v) = \langle F(x, u, t), u - v \rangle - \alpha \phi(u, v) \leq -\alpha \phi(u, v)$$

for all $v \in U$. Furthermore, we have

$$G_\alpha(u) = \max_{v \in U} \Phi_\alpha(u, v) \leq -\alpha\phi(u, v_\alpha)$$

which contradicts the nonnegativity property of $G_\alpha(u)$, unless $G_\alpha(u) = 0$ and $u = v_\alpha(u)$. ■

The above result assures that a broad class of gap functions for differential variational inequalities may be defined.

6.5.2 D-gap Equivalent Optimal Control Problem

Note that the preceding definition of the gap function does not ensure that $G_\alpha(u)$ is in general differentiable, a limitation we would like to overcome. To that end, let us introduce the so-called a *D-gap function*, which is based on the primitive gap functions G_α and G_β introduced above and has the form

$$\psi_{\alpha\beta}(u) = G_\alpha(u) - G_\beta(u) \tag{6.72}$$

for $0 < \alpha < \beta$. While $G_\alpha(u)$ is not differentiable in general, $\psi_{\alpha\beta}(u)$ is Gateaux-differentiable, as we demonstrate subsequently. To show that $\psi_{\alpha\beta}(u)$ is a gap function, we only need to show the essential nonnegativity property holds. We continue to invoke assumptions A1, A2, A3, and A4; hence, by virtue of strong convexity of $\phi(u)$, we have

$$\begin{aligned}
\psi_{\alpha\beta}(u) &= G_\alpha(u) - G_\beta(u) \\
&= \Phi_\alpha(u, v_\alpha) - \Phi_\beta\left(u, v_\beta\right) \\
&\geq \Phi_\alpha\left(u, v_\beta\right) - \Phi_\beta\left(u, v_\beta\right) \\
&= \left\langle F(x, u, t), u - v_\beta \right\rangle - \alpha\phi\left(u, v_\beta\right) - \left\langle F(x, u, t), u - v_\beta \right\rangle + \beta\phi\left(u, v_\beta\right) \\
&= (\beta - \alpha)\,\phi\left(u, v_\beta\right)
\end{aligned}$$

This demonstrates $\psi_{\alpha\beta}(u) \geq 0$, and, of course, $G_\alpha(u) \geq G_\beta(u)$ for all $u \in U$. So (6.72) does in fact define a gap function.

We will employ, as our D-gap function, the gap function Fukushima (1992) has named the *regularized gap function* for finite-dimensional spaces and Konnov et al. (2002) have extended to Hilbert spaces. In particular, we introduce as a generator of the regularized gap function the following

$$\phi(u, v) = \frac{1}{2}\|v - u\|^2 \tag{6.73}$$

which satisfies the relevant assumptions on $\phi(\cdot)$, especially that of strong convexity with modulus $\alpha > 0$. From (6.64) and (6.73) we get

$$\Phi_\alpha(u, v) = \langle F(x, u, t), u - v\rangle - \frac{\alpha}{2}\|v - u\|^2$$

The corresponding D-gap function becomes

$$\psi_{\alpha\beta}(u) = G_\alpha(u) - G_\beta(u) = \max_{v\in U}\Phi_\alpha(u, v) - \max_{v\in U}\Phi_\beta(u, v)$$

or, alternatively

$$\psi_{\alpha\beta}(u) = \langle F(x, u, t), v_\beta(u) - v_\alpha(u)\rangle - \frac{\alpha}{2}\|v_\alpha(u) - u\|^2 + \frac{\beta}{2}\|v_\beta(u) - u\|^2 \tag{6.74}$$

where

$$v_\alpha(u) = \arg\max_{v\in U}\Phi_\alpha(u, v) \tag{6.75}$$

$$v_\beta(u) = \arg\max_{v\in U}\Phi_\beta(u, v) \tag{6.76}$$

Furthermore, it should be noted that, for a fixed $u \in U$, the maximization problem (6.75) is equivalent to the following:

$$v_\alpha(u) = \arg\min_{v\in U}\left\|v - \left(u - \frac{1}{\alpha}F(x, u, t)\right)\right\|^2$$

which may be rewritten in the form of a fixed-point problem involving a projection operator, namely

$$v_\alpha(u) = P_U\left[u - \frac{1}{\alpha}F(x, u, t)\right] \tag{6.77}$$

as observed in Fukushima (1992) for finite dimensions and Konnov et al. (2002) for infinite dimensions.

Let us now consider the D-gap function in the context of an unconstrained differential variational inequality without terminal-time constraints. That is, we stipulate the following

$$\begin{gathered} U = \left(L^2[t_0, t_f]\right)^m \\ x(t_f) \text{ free} \end{gathered}$$

Thus, in terms of our notation, we are interested in restating the problem

$$DVI(F, f, \left(L^2[t_0, t_f]\right)^m, x(t_f) \text{ free}, x_0, t_0, t_f) \tag{6.78}$$

in terms of the D-gap function (6.74). That restatement yields the following equivalent optimal control problem:

$$\min \ \psi_{\alpha\beta}(u) = \int_{t_0}^{t_f} F_0(x,u,t)\,dt \tag{6.79}$$

subject to

$$\frac{dx}{dt} = f(x,u,t) \tag{6.80}$$

$$x(t_0) = x_0 \tag{6.81}$$

where

$$F_0(x,u,t) \equiv F(x,u,t)\left[v_\beta(u) - v_\alpha(u)\right] - \frac{\alpha}{2}\left[v_\alpha(u) - u\right]^2 + \frac{\beta}{2}\left[v_\beta(u) - u\right]^2 \tag{6.82}$$

The criterion integrand (6.82) is determined by observing that

$$\begin{aligned}\psi_{\alpha\beta}(u) &= \left\langle F(x,u,t), v_\beta(u) - v_\alpha(u)\right\rangle - \frac{\alpha}{2}\left\|v_\alpha(u) - u\right\|^2 + \frac{\beta}{2}\left\|v_\beta(u) - u\right\|^2 \\ &= \int_{t_0}^{t_f}\left\{F(x,u,t)\left[v_\beta(u) - v_\alpha(u)\right] - \frac{\alpha}{2}\left[v_\alpha(u) - u\right]^2 + \frac{\beta}{2}\left[v_\beta(u) - u\right]^2\right\}dt\end{aligned}$$

A more general differential variational inequality than (6.78) could be considered without complications other than increased notational complexity. The problem (6.79), (6.80), and (6.81) is a Bolza-form optimal control problem; it is unusual only in that the objective functional involves the maximizers of subproblems defined by (6.75) and (6.76), namely $v_\alpha(u)$ and $v_\beta(u)$.

Now we are interested in the gradient of the objective functional $\psi_{\alpha\beta}(u)$, which is equivalent to the gradient of the corresponding Hamiltonian, owing to the particular function spaces we have elected in this chapter. In particular, the Hamiltonian for problem (6.79), (6.80), and (6.81) is

$$\begin{aligned}H_{\alpha\beta}(x,u,\lambda,t) = {} & F(x,u,t)\left[v_\beta(u) - v_\alpha(u)\right] - \frac{\alpha}{2}\left[v_\alpha(u) - u\right]^2 + \frac{\beta}{2}\left[v_\beta(u) - u\right]^2 \\ & + \lambda f(x,u,t)\end{aligned} \tag{6.83}$$

To obtain the gradient of $\psi_{\alpha\beta}(u)$, we need to carefully consider the role of $v_\alpha(\cdot)$ and $v_\beta(\cdot)$ which are maximizers defined by (6.75) and (6.76). In particular, $v_\alpha(\cdot)$ and $v_\beta(\cdot)$ are unique by the strong concavity of $\Phi_\alpha(u,v)$ and $\Phi_\beta(u,v)$ in v and the convexity of the set U. To continue our analysis, we employ, without proof, the following lemma from Pshenichnyi (1971):

Lemma 6.2. *Gradient and G-derivatives. Let V be an abstract Hilbert space, $U \subseteq V$ and $h : V \times U \to \Re$ a mapping whose gradient $\nabla_u h(u,v)$ exists everywhere on U and is continuous on $V \times U$. Define two functions as follows:*

$$w(u) = \max_{v \in U} h(u, v)$$
$$z(u) = \{v \in U : w(u) = h(u, v)\}$$

Then the G-derivative of $w(u)$ and the G-derivative of $h(u, v)$ in the direction ρ are related according to

$$\delta w(u, \rho) = \max_{v \in z(u)} \delta h(u, \rho; v, \rho)$$

Furthermore, if $z(u)$ is a singleton for all $u \in V$ and z is a continuous function on V, then w is continuously differentiable and its gradient is given by

$$\nabla w(u) = \nabla \max_{v \in U} h(u, v) = \nabla_u h(u, z(u)).$$

Now we are in a position to articulate the gradient of the objective functional $\psi_{\alpha\beta}(u)$:

Theorem 6.9. *Gradient of D-gap function. Suppose $F(x, u, t)$ is Lipschitz continuous on every bounded subset of $\left(L^2\left[t_0, t_f\right]\right)^m$. Then $\psi_{\alpha\beta}(u)$ is continuously differentiable in the sense of Gateaux and*

$$\begin{aligned}\nabla \psi_{\alpha\beta}(u) &= \frac{\partial}{\partial u} H_{\alpha\beta}(x, u, \lambda, t) \\ &= \frac{\partial F(x, u, t)}{\partial u}\left[v_\beta(u) - v_\alpha(u)\right] \\ &\quad + \alpha\left[v_\alpha(u) - u\right] - \beta\left[v_\beta(u) - u\right] + \lambda \frac{\partial f(x, u, t)}{\partial u}\end{aligned}$$

Proof. Rewrite the objective functional as

$$\begin{aligned}\psi_{\alpha\beta}(u) &= G_\alpha(u) - G_\beta(u) \\ &= \max_{v \in U}\left\{\int_{t_0}^{t_f} F(x, u, t)[u - v] - \frac{\alpha}{2}[v - u]^2 dt\right\} \\ &\quad - \max_{v \in U}\left\{\int_{t_0}^{t_f} F(x, u, t)[u - v] - \frac{\beta}{2}[v - u]^2 dt\right\}\end{aligned}$$

Let us define

$$g_\alpha(u, v) = F(x, u, t)[u - v] - \frac{\alpha}{2}[v - u]^2$$
$$g_\beta(u, v) = F(x, u, t)[u - v] - \frac{\beta}{2}[v - u]^2$$

Then, by Lemma 6.2 we have

$$\delta G_\alpha(u, \rho) = \int_{t_0}^{t_f}\left\{\frac{\partial g_\alpha(u, v_\alpha)}{\partial x} y + \frac{\partial g_\alpha(u, v_\alpha)}{\partial u}\rho\right\} dt$$

$$\delta G_\beta(u,\rho)=\int_{t_0}^{t_f}\left\{\frac{\partial g_\beta\left(u,v_\beta\right)}{\partial x}y+\frac{\partial g_\beta\left(u,v_\beta\right)}{\partial u}\rho\right\}dt$$

so that

$$\delta\psi_{\alpha\beta}(u,\rho)=\int_{t_0}^{t_f}\left\{\frac{\partial g}{\partial x}y+\frac{\partial g}{\partial u}\rho\right\}dt \tag{6.84}$$

where we for simplicity of notation we write

$$g(u)=g_\alpha(u,v_\alpha)-g_\beta\left(u,v_\beta\right)$$

and $y=\delta x$ is a variation in x which implicitly depends on ρ. Furthermore, by definition

$$x(t)=x_0+\int_{t_0}^{t}f(x,u,y)\,dt$$

Also, we know from our analysis of continuous-time optimal control problems in Chapter 4 that

$$y=\int_{t_0}^{t}\left[\frac{\partial f}{\partial x}y+\frac{\partial f}{\partial u}\rho\right]dt$$

We introduce the adjoint vector defined by the final-value problem

$$-\frac{d\lambda}{dt}=\left(\frac{\partial f}{\partial x}\right)^T\lambda+\left(\frac{\partial g}{\partial x}\right)^T \tag{6.85}$$

$$\lambda\left(t_f\right)=0 \tag{6.86}$$

so that (6.84) becomes

$$\delta\psi_{\alpha\beta}(u,\rho)=\int_{t_0}^{t_f}\left\{\left[-\left(\frac{d\lambda}{dt}\right)^T-\lambda^T\frac{\partial f}{\partial x}\right]y+\frac{\partial g}{\partial u}\rho\right\}dt$$

Noting that $y(t_0)=0$ and $\lambda\left(t_f\right)=0$, the by now familiar step of integration by parts yields

$$\begin{aligned}\int_{t_0}^{t_f}-\left(\frac{d\lambda}{dt}\right)^T ydt&=\int_{t_0}^{t_f}\lambda^T\frac{dy}{dt}dt\\&=\int_{t_0}^{t_f}\lambda^T\left[\frac{\partial f}{\partial x}y+\frac{\partial f}{\partial u}\rho\right]dt\end{aligned}$$

It follows that

$$\delta\psi_{\alpha\beta}(u,\rho)=\int_{t_0}^{t_f}\left\{\lambda^T\left[\frac{\partial f}{\partial x}y+\frac{\partial f}{\partial u}\rho\right]-\lambda^T\frac{\partial f}{\partial x}y+\frac{\partial g}{\partial u}\rho\right\}dt$$

$$= \int_{t_0}^{t_f} \left\{ \lambda^T \frac{\partial f}{\partial u} + \frac{\partial g}{\partial u} \right\} \rho dt$$
$$= \left\langle \lambda^T \frac{\partial f}{\partial u} + \frac{\partial g}{\partial u}, \rho \right\rangle$$

Therefore, the gradient of $\psi_{\alpha\beta}(u)$ becomes

$$\nabla \psi_{\alpha\beta}(u) = \lambda^T \frac{\partial f}{\partial u} + \frac{\partial g}{\partial u}$$
$$= \nabla_u H_{\alpha\beta}(x, u, \lambda, t)$$

Furthermore, we note that

$$\nabla \psi_{\alpha\beta}(u) = \nabla_u H_{\alpha\beta}(x, u, \lambda, t)$$
$$= \lambda \frac{\partial f(x, u, t)}{\partial u} + \frac{\partial F(x, u, t)}{\partial u} \left[v_\beta(u) - v_\alpha(u) \right]$$
$$+ \alpha \left[v_\alpha(u) - u \right] - \beta \left[v_\beta(u) - u \right]$$

Also, we note that

$$-\frac{d\lambda}{dt} = \nabla_x H_{\alpha\beta}(x, u, \lambda, t)$$

which is recognized as the adjoint equation. ■

To solve an extremal problem with fixed initial time, fixed terminal time, and free terminal state using the D-gap function in Hilbert space, one needs certain additional information. In particular, the terminal time constraint

$$\Psi \left[x(t_f), t_f \right] = 0 \tag{6.87}$$

and the final value of the adjoint vector

$$\lambda(t_f) = \mu^T \frac{\partial \Psi \left[x(t_f), t_f \right]}{\partial x}$$

where μ is the Lagrange multiplier for (6.87).

6.5.3 Numerical Example

Let us consider the following example of *DVI(F, f, Ψ, U, x₀, t₀, t_f)* from Friesz and Mookherjee (2006), again involving three controls and two states:

$$u \in \left(L^2[0, 1] \right)^3$$

$$x \in \left(\mathcal{H}^1[0, 1] \right)^2$$

$$x\left(t_0\right)=\begin{pmatrix}1\\0.7\end{pmatrix}$$

$$\left[t_0,t_f\right]=[0,5]$$

$$F(x,u)=\begin{pmatrix}F_1(x,u)\\F_2(x,u)\\F_3(x,u)\end{pmatrix}=\begin{pmatrix}x_1^2-u_1\left(t\right)+u_2\left(t\right)\\x_2-u_2^2\left(t\right)-u_3\left(t\right)\\\frac{1}{10}x_2^2-u_3^2\left(t\right)\end{pmatrix}$$

$$f\left(x,u\right)=\begin{pmatrix}f_1(x,u)\\f_2(x,u)\end{pmatrix}=\begin{pmatrix}\frac{1}{5}x_1\left(t\right)+\frac{1}{2}u_1\left(t\right)+\frac{3}{10}u_2\left(t\right)\\\frac{1}{4}x_2\left(t\right)+\frac{1}{2}u_2\left(t\right)-\frac{1}{5}u_3\left(t\right)\end{pmatrix}$$

$$U=\{u:0.2\leq u_1\leq 1;\ 0.2\leq u_2\leq 1.2;0.2\leq u_3\leq 1.3\}$$

Results using a projected gradient/D-gap function in Hilbert space are presented in Figures 6.2 and 6.3. This example was solved using MATLAB on a PC with an Intel Xeon 3.06 GHz CPU and 3.37 GB RAM. The computation time is less than 20 seconds. The algorithm converged in 11 iterations involving 22 subproblems with a gap size less than 10^{-10}.

6.6 Differential Variational Inequalities with Time Shifts

We are now ready to consider the formulation of differential inequalities with state-dependent time shifts. In particular, we retain as much of our prior notation as possible and consider:

$$\left.\begin{array}{c}u^*\in U\\\left\langle F\left(x\left(u^*,u_\tau^*\right),u^*,u_\tau^*,t\right),u-u^*\right\rangle\geq 0\quad\forall u\in U\end{array}\right\} \tag{6.88}$$

where

$$x\left(u,u_\tau,t\right)=$$

$$\arg\left\{\frac{dx}{dt}=f\left(x,u,u_\tau,t\right),\ x\left(t_0\right)=x_0,u\in U,\Psi\left[x\left(t_f\right),t_f\right]=0\right\}\in\left(\mathcal{H}^1\left[t_0,t_f\right]\right)^n \tag{6.89}$$

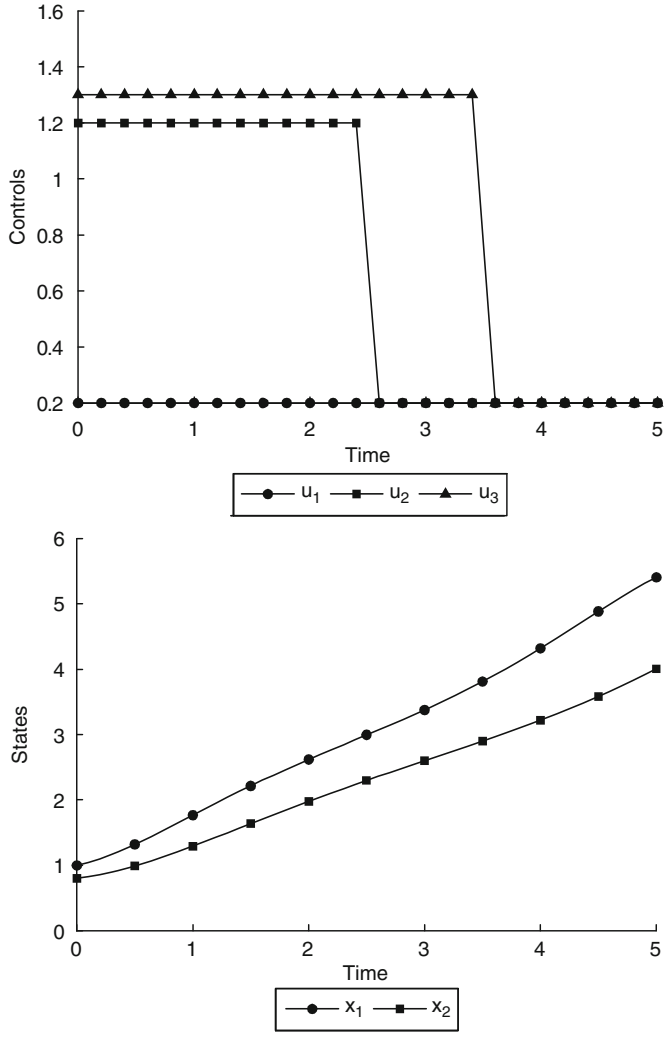


Fig. 6.2 Result by gap function (gap $< 10^{-10}$, $\alpha = 0.5$, $\beta = 2$)

for

$$\left[t_0, t_f\right] \subseteq \Re_+^1$$

Furthermore $u_\tau(t)$ is a shorthand for the shifted control vector

$$u_\tau(t) = \begin{pmatrix} u_1(t - \tau_1(x_1)) \\ \vdots \\ u_m(t - \tau_m(x_m)) \end{pmatrix}$$

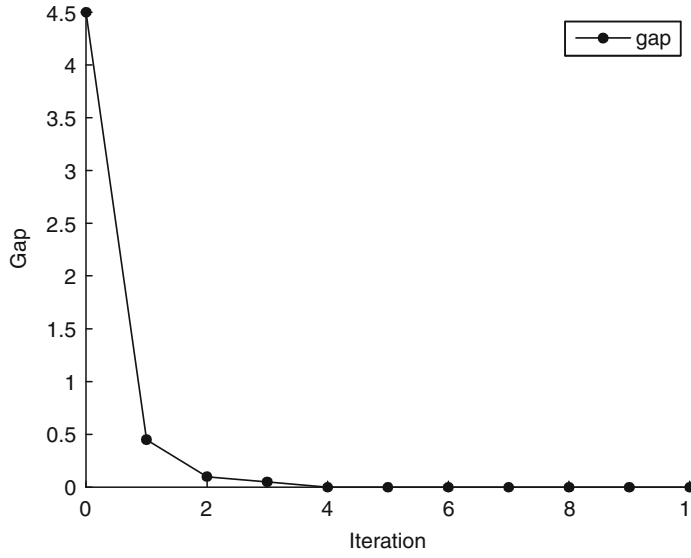


Fig. 6.3 Convergence of the descent algorithm, which is terminated with the gap less than 10^{-10}

where

$$\begin{aligned} \tau_i &: \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \longrightarrow \Re^1_+ \\ \tau &= (\tau_i : i \in [1, m]) \end{aligned}$$

for each $i \in [1, m]$. The other relevant mappings are

$$\begin{aligned} f &: \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, \tau\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re^1_+ \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n \\ \Psi &: \Re^n \times \Re^1_+ \longrightarrow \Re^r \\ u &\in U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m \\ u_\tau &\in \left(L^2\left[t_0, t_f\right]\right)^m \end{aligned}$$

Of course, in the above $\left(L^2\left[t_0, t_f\right]\right)^m$ is the m-fold product of the space of square-integrable functions $L^2\left[t_0, t_f\right]$, while $\left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ is the n-fold product of the Sobolev space $\mathcal{H}^1\left[t_0, t_f\right]$. We refer to (6.88) as a differential variational inequality with time shifts, abbreviated *DVI(F, f, Ψ, U, x_0, t_0, t_f; τ)*.

6.6.1 Necessary Conditions

To develop necessary conditions for *DVI(F, f, Ψ, U, x_0, t_0, t_f, τ)*, we will rely on the following notion of regularity:

Definition 6.5. *Regularity of DVI(F, f, Ψ, U, x_0, t_0, t_f, τ). We call DVI(F, f, Ψ, U, x_0, t_0, t_f, τ) regular if*

R1. $u \in U \subseteq \left(L^2\left[t_0, t_f\right]\right)^m$;
R2. $u_\tau \in U \in \left(L^2\left[t_0, t_f\right]\right)^m$;
R3. $x \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$
R4. $x\left(u, u_\tau, t\right) : \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re_+^1 \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n$ *exists for all admissible u and is unique, strongly continuous, and G-differentiable with respect to u and* u_τ;
R5. τ *is continuously differentiable with respect to x;*
R6. $\Psi : \Re^n \times \Re_+^1 \longrightarrow \Re^r$ *is continuously differentiable with respect to x and t;*
R7. $F(x, u, u_\tau, t) : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re_+^1 \longrightarrow L^2\left[t_0, t_f\right]$ *is continuous with respect to x, u, and* u_τ;
R8. $f : \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^n \times \left(L^2\left[t_0, t_f\right]\right)^m \times \left(L^2\left[t_0, t_f\right]\right)^m \times \Re_+^1 \longrightarrow \left(L^2\left[t_0, t_f\right]\right)^n$ *is continuously differentiable with respect to x, u, and* u_τ;
R9. $x_0 \in \Re^n$, $t_0 \in \Re_+^1$, *and* $t_f \in \Re_{++}^1$ *are known and fixed;*
R10. $U \subset \left(L^2\left[t_0, t_f\right]\right)^m$ *is convex; and*
R11. there is a constant dual vector $\gamma \in \Re^r$ *for the terminal constraints* $\Psi\left[x\left(t_f\right), t_f\right] = 0$.

We next note that (6.88) may be restated as the following optimal control problem

$$\min \gamma^T \Psi\left[x\left(t_f\right), t_f\right] + \int_{t_0}^{t_f} \left[F\left(x^*, u^*, u_\tau^*, t\right)\right]^T u dt \tag{6.90}$$

subject to

$$\frac{dx}{dt} = f\left(x, u, u_\tau, t\right) \quad x\left(t_0\right) = x_0 \tag{6.91}$$

$$u \in U \tag{6.92}$$

where $x^* = x\left(u^*, u_\tau^*, t\right)$ is the optimal state vector and $\gamma \in \Re^r$ is the vector of dual variables for the terminal constraints $\Psi\left[x\left(t_f\right), t_f\right] = 0$. We point out that this optimal control problem is a mathematical abstraction and of no use for computation, since its criterion depends on knowledge of the variational inequality solution u^*. In what follows we will need the Hamiltonian for (6.90), (6.91), and (6.92), namely

$$H_2\left(x, u, u_\tau, \lambda, t\right) = \left[F\left(x^*, u^*, u_\tau^*, t\right)\right]^T u + \lambda^T f\left(x, u, u_\tau, t\right) \tag{6.93}$$

where $\lambda\left(t\right)$ is the adjoint vector that solves the adjoint equations and transversality conditions for given state variables and controls. It is now a relatively easy matter to derive the necessary conditions stated in the following theorem:

Theorem 6.10. *Necessary conditions for DVI(F, f,* Ψ, *U,* x_0, t_0, t_f, τ*). When* x_0, t_0, *and* t_f *are fixed and regularity in the sense of Definition 6.5 holds, necessary conditions for* $u^* \in U$, *a solution of DVI(F, f,* Ψ, *U,* x_0, t_0, t_f, τ*), are:*

1. *the variational inequality principle:*

$$\sum_{i=1}^{m} \frac{\partial H_2\left(x^*, u^*, u_\tau^*, \lambda^*, t\right)}{\partial u_i}\left(u_i - u_i^*\right) \geq 0 \quad \forall u \in U$$

where

$$H_2\left(x, u, u_\tau, \lambda, t\right) = \left[F\left(x^*, u^*, u_\tau^*, t\right)\right]^T u + \lambda^T f\left(x, u, u_\tau, t\right)$$

$$\frac{\partial H_2}{\partial u_i} = F_i\left(x^*, u^*, u_\tau^*, t\right) + \sum_{j=1}^{m} \lambda_j \frac{\partial f_i\left(x^*, u^*, u_\tau^*, t\right)}{\partial u_i}$$
$$if\ t \in \left[t_f - \tau_i\left(x\left(t_f\right)\right), t_f\right]$$

$$\frac{\partial H_2}{\partial u_i} = F_i\left(x^*, u^*, u_\tau^*, t\right) + \sum_{j=1}^{m} \lambda_j \frac{\partial f_j\left(x^*, u^*, u_\tau^*, t\right)}{\partial u_i}$$
$$+ \left[\lambda_j \frac{\partial f_j\left(x^*, u^*, u_\tau^*, t\right)}{\partial\left(u_\tau\right)_i} \frac{1}{1 - \sum_{j=1}^{m} \frac{\partial \tau_i\left(x^*\right)}{\partial x_j} f_j\left(x^*, u^*, u_\tau^*, t\right)}\right]_{s_i}$$
$$if\ \ t \in \left[t_0, t_f - \tau_i\left(x^*\left(t_f\right)\right)\right]$$

and each s_i solves the fixed point problem

$$t = s_i + \tau_i\left[x\left(t\right)\right]$$

2. *the state dynamics:*

$$\frac{dx^*}{dt} = f\left(x^*, u^*, u_\tau^*, t\right)$$
$$x^*\left(t_0\right) = x_0$$

3. *the adjoint dynamics:*

$$(-1)\frac{d\lambda^*}{dt} = \nabla_x\left(\lambda^*\right)^T f\left(x^*, u^*, u_\tau^*, t\right)$$
$$\lambda^*\left(t_f\right) = \gamma^T \frac{\partial \Psi\left[x^*\left(t_f\right), t_f\right]}{\partial x}$$

where $\gamma \in \Re^r$ is the vector of dual variables for the terminal constraint

$$\Psi\left[x\left(t_f\right), t_f\right] = 0$$

Proof. *DVI(F, f, Ψ, U, x_0, t_0, t_f, τ)* is equivalent to the optimal control problem (6.90), (6.91), and (6.92), with Hamiltonian (6.93). By virtue of regularity, we may employ the necessary conditions for optimal control problems with state-dependent time shifts from Chapter 4; the relevant differential variational inequality necessary conditions follow immediately. ■

6.6.2 Fixed-Point Formulation and Algorithm

There is a fixed-point form of *DVI(F, f, Ψ, U, x_0, t_0, t_f, τ)*. In particular we state and prove the following result:

Theorem 6.11. *Fixed-point formulation of DVI(F, f, Ψ, U, x_0, t_0, t_f, τ). When regularity in the sense of Definition 6.5 holds, any fixed-point of*

$$u = P_U \left[u - \alpha F \left(x \left(u, u_\tau, t \right), u, u_\tau, t \right) \right]$$

must be a solution of DVI(F, f, Ψ, U, x_0, t_0, t_f, τ), where $P_U[.]$ is the minimum norm projection onto $U \subseteq \left(L^2 \left[t_0, \tau \right] \right)^m$ and $\alpha \in \Re^1_{++}$.

Proof. The proof is very similar to the case of no time shifts; however, in the interest of gaining familiarity with time-shifted problems, it is worth providing a detailed exposition. In particular, we note the fixed-point problem considered requires that

$$u = \arg\min_v \left\{ \frac{1}{2} \left\| u - \alpha F \left(x \left(u, u_\tau, t \right), u, u_\tau, t \right) - v \right\|^2 : v \in U \right\} \tag{6.94}$$

where $\alpha \in \Re^1_{++}$ is any strictly positive real number. That is, we seek the solution of the optimal control problem

$$\min_v \gamma^T \Psi \left[x \left(t_f \right), t_f \right] + \int_{t_0}^{t_f} \frac{1}{2} \left[u - \alpha F \left(x, u, u_\tau, t \right) - v \right]^2 dt \tag{6.95}$$

subject to

$$\frac{dx}{dt} = f \left(x, v, v_\tau, t \right); \; x \left(t_0 \right) = x_0 \tag{6.96}$$

$$v \in U \tag{6.97}$$

where u and u_τ are treated as fixed vectors. Any solution of optimal control problem (6.95), (6.96), and (6.97) must satisfy the necessary conditions of Theorem 6.4. In particular, we must have

$$\left[\nabla_v H_3 \left(x^*, v^*, v_\tau^*, \eta^*, t \right) \right]^T \left(v - v^* \right) \geq 0 \qquad \forall v \in U \tag{6.98}$$

where

$$H_3(x, v, v_\tau, \eta, t) = \frac{1}{2}[u - \alpha F(x, u, u_\tau, t) - v]^2 + \eta^T f(x, v, v_\tau, t)$$

and for given x and v

$$\eta = \arg\left\{(-1)\frac{d\eta}{dt} = \nabla_x H_3(x, v, v_\tau, \eta, t),\ \eta(t_f) = \gamma^T \frac{\partial \Psi[x(t_f), t_f]}{\partial x(t_f)}\right\}$$

Note that

$$\nabla_v H_3(x, v, v_\tau, \eta, t) = -u + \alpha F(x, u, u_\tau, t) + v + \nabla_v \eta^T f(x, v, v_\tau, t)$$

Because $u = v$ by virtue of (6.94) we have

$$\nabla_u H_3(x, v, v_\tau, \eta, t) = \alpha F(x, u, u_\tau, t) + \nabla_u \eta^T f(x, u, u_\tau, t) \tag{6.99}$$

Now if we set $\lambda = \dfrac{\eta}{\alpha}$; we have

$$\left[F(x^*, u^*, u_\tau^*, t) + \nabla_u(\lambda^*)^T f(x^*, u^*, t)\right]^T (u - u^*) \geq 0 \qquad \forall v \in U$$

which is identical to the finite-dimensional variational inequality principle of Theorem 6.10. The other optimality conditions are also identical. This completes the proof. ■

Naturally there is an associated fixed-point algorithm based on the iterative scheme

$$u^{k+1} = P_U\left[u^k - \alpha F\left(x\left(u^k, u_\tau^k\right), u^k, u_\tau^k, t\right)\right]$$

The detailed structure of the fixed-point algorithm is:

Step 0. Initialization. Identify an initial feasible solution $u^0 \in U$ and set $k = 0$.

Step 1. Solve optimal control problem. Solve the following optimal control problem:

$$\min_v J^k(v) = \gamma^T \Psi[x(t_f), t_f] + \int_{t_0}^{t_f} \frac{1}{2}\left[u^k - \alpha F\left(x^k, u^k, u_\tau^k, t\right) - v\right]^2 dt \tag{6.100}$$

$$\text{subject to } \frac{dx}{dt} = f(x, v, v_\tau, t) \quad x(t_0) = x_0 \tag{6.101}$$

$$v \in U \tag{6.102}$$

Call the solution u^{k+1}.

Step 2. Stopping test. If $\left\|u^{k+1} - u^k\right\| \leq \varepsilon$ where $\varepsilon \in \Re^1_{++}$ is a preset tolerance, stop and declare $u^* \approx u^{k+1}$. Otherwise set $k = k + 1$ and go to Step 1.

Note that, if time shifts do appear in the principal operator F, but not in the dynamics or constraints of the time-shifted differential variational inequality whose solution is sought, the fixed-point subproblems formed by (6.100), (6.101), and (6.102) have the appealing property that they are conventional (not time-shifted). This means that the subproblems may be solved by conventional methods. In particular, for the special circumstance we have mentioned, gradient projection will be a convergent algorithm for the subproblems, provided appropriate regularity conditions for convergence without time shifts are met.

6.6.3 Time-Shifted Numerical Examples

In this section we provide three related numerical examples of *DVI(F, f, Ψ, U, x_0, t_0, t_f, τ)* with time shifts to illustrate a fixed-point, descent-in-Hilbert-space solution scheme. The computations were performed using Matlab 6.5 on a Pentium 4 processor desktop computer with 1 GB RAM. The three examples differ from one another according to what type of time shift is employed. In particular, we consider both fixed and state-dependent time shifts as well as the degenerate case of no of time shifts. The run times for these examples were found to be less than 1 minute for the computing hardware described above.

Example 1 (State-Dependent Time Shifts)

Consider a version of *DVI(F, f, Ψ, U, x_0, t_0, t_f, τ)* involving three controls and two states:

$$u \in \left(L^2\left[0,1\right]\right)^3 \quad x \in \left(\mathcal{H}\left[0,1\right]\right)^2 \quad x\left(0\right) = \begin{pmatrix} 1 \\ 0.7 \end{pmatrix} \quad t_f = 5$$

$$F(x,u,u_\tau) = \begin{pmatrix} F_1(x,u,u_\tau) \\ F_2(x,u,u_\tau) \\ F_3(x,u,u_\tau) \end{pmatrix} = \begin{pmatrix} x_1^2 - u_1\left(t+\tau_1\left(x\right)\right) + u_2\left(t+\tau_2\left(x\right)\right) \\ x_2 - u_2^2\left(t+\tau_2\left(x\right)\right) - u_3\left(t\right) \\ \frac{1}{10}x_2^2 - u_3^2\left(t+\tau_3\left(x\right)\right) \end{pmatrix}$$

$$f\left(x,u\right) = \begin{pmatrix} f_1(x,u,u_\tau) \\ f_2(x,u,u_\tau) \end{pmatrix} = \begin{pmatrix} \frac{1}{5}x_1\left(t\right) + \frac{1}{2}u_1\left(t\right) + \frac{3}{10}u_2\left(t+\tau_2\left(x\right)\right) \\ \frac{1}{4}x_2\left(t\right) + \frac{1}{2}u_2\left(t\right) - \frac{1}{5}u_3\left(t+\tau_3\left(x\right)\right) \end{pmatrix}$$

$$U = \{u : 1 \geq u_1 \geq 0.2,\ 1.2 \geq u_2 \geq 0.2,\ 1.3 \geq u_3 \geq 0.2\}$$

Note that the righthand side of the state dynamics has shifted controls; the shifted control vectors obey

$$u_\tau = \begin{bmatrix} u_1\left(t+\tau_1\left(x\right)\right) \\ u_2\left(t+\tau_2\left(x\right)\right) \\ u_3\left(t+\tau_3\left(x\right)\right) \end{bmatrix}$$

where

$$\begin{aligned} \tau_1\left(x\left(t\right)\right) &= \frac{x_1\left(t\right)}{k_1} \\ \tau_2\left(x\left(t\right)\right) &= \frac{x_2\left(t\right)}{k_2} \\ \tau_3\left(x\left(t\right)\right) &= \frac{0.7\cdot x_1\left(t\right)+0.3\cdot x_2\left(t\right)}{k_3} \end{aligned}$$

and

$$\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} 80 \\ 85 \\ 90 \end{pmatrix}$$

We choose the fixed-point parameter to be $\alpha = 0.05$. A fifth-power polynomial was used to express the controls, adjoint variables and state variables as continuous functions of time. Also, the nominal decision time interval is $[0, 5]$. The stopping tolerances for both fixed-point and descent iterations were set at $\varepsilon = 10^{-2}$. The combined fixed-point, descent-in-Hilbert-space algorithm converged after 17 fixed-point iterations; each of the descent subproblems converged in 10 or fewer iterations. We forgo the detailed symbolic statement of this example and, instead, provide numerical results in graphical form. Figure 6.4 shows the controls u^* and the states x^* plotted against time.

Example 2 (Fixed Time Shifts)

Next we modify Example 1 so that the shifts do not depend on the states; instead they are fixed. In particular we assume

$$\begin{aligned} \tau_1 &= \frac{3.5}{k_1} \\ \tau_2 &= \frac{2}{k_2} \\ \tau_3 &= \frac{3}{k_3} \end{aligned}$$

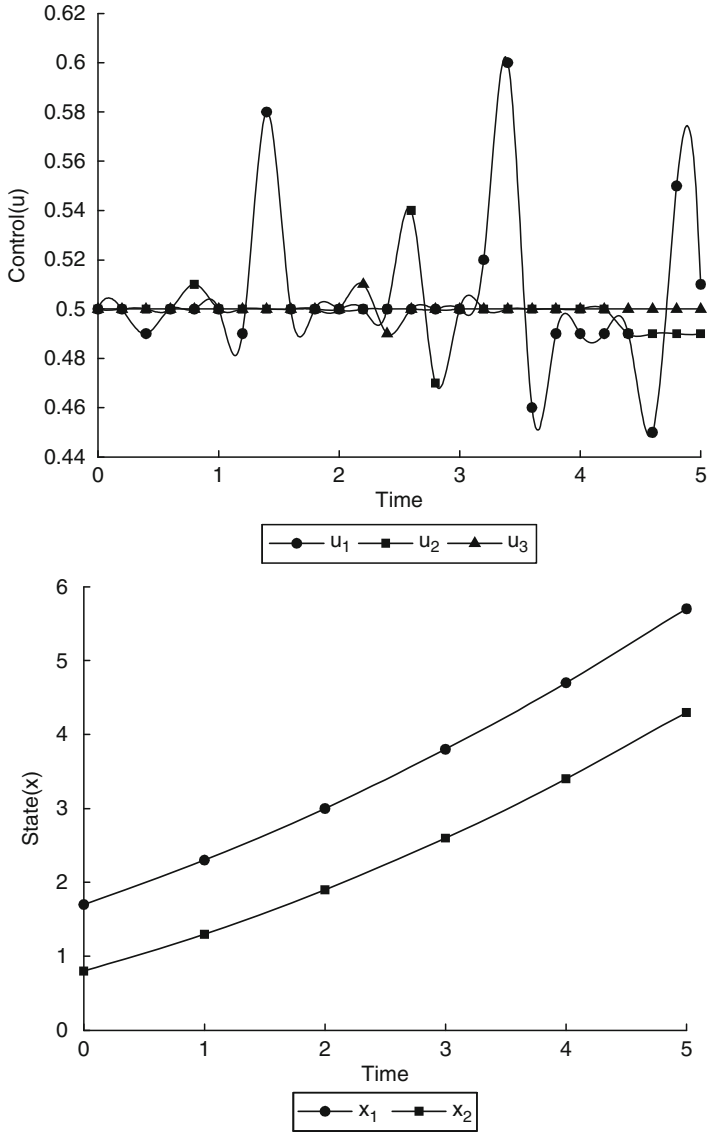


Fig. 6.4 u^* vs. time and x^* vs. time with state-dependent time shifts

and keep all other parameters the same. The solution obtained is shown in Figure 6.5. In this case, the combined fixed-point, descent-in-Hilbert-space algorithm converged in 15 fixed-point iterations; each of the descent subproblems converged in 12 or fewer iterations.

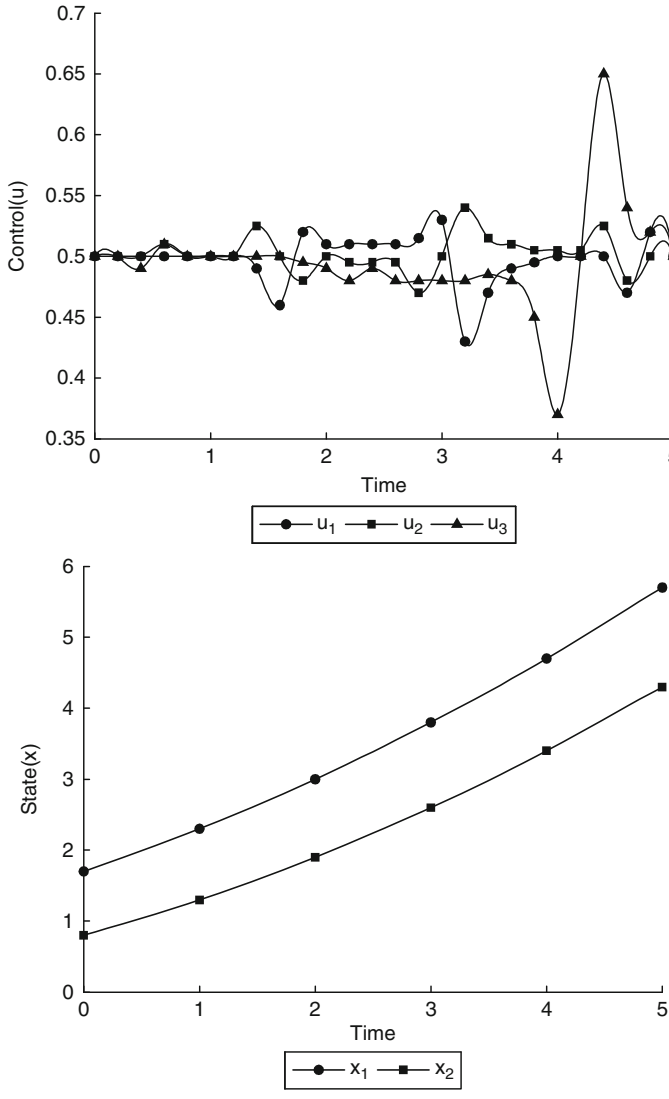


Fig. 6.5 u^* vs. time and x^* vs. time with fixed time shifts

Example 3 (Degenerate Case: No Time Shifts)

Next we modify Example 1 so that there are no time shifts; that is

$$\tau_1 = \tau_2 = \tau_3 = 0$$

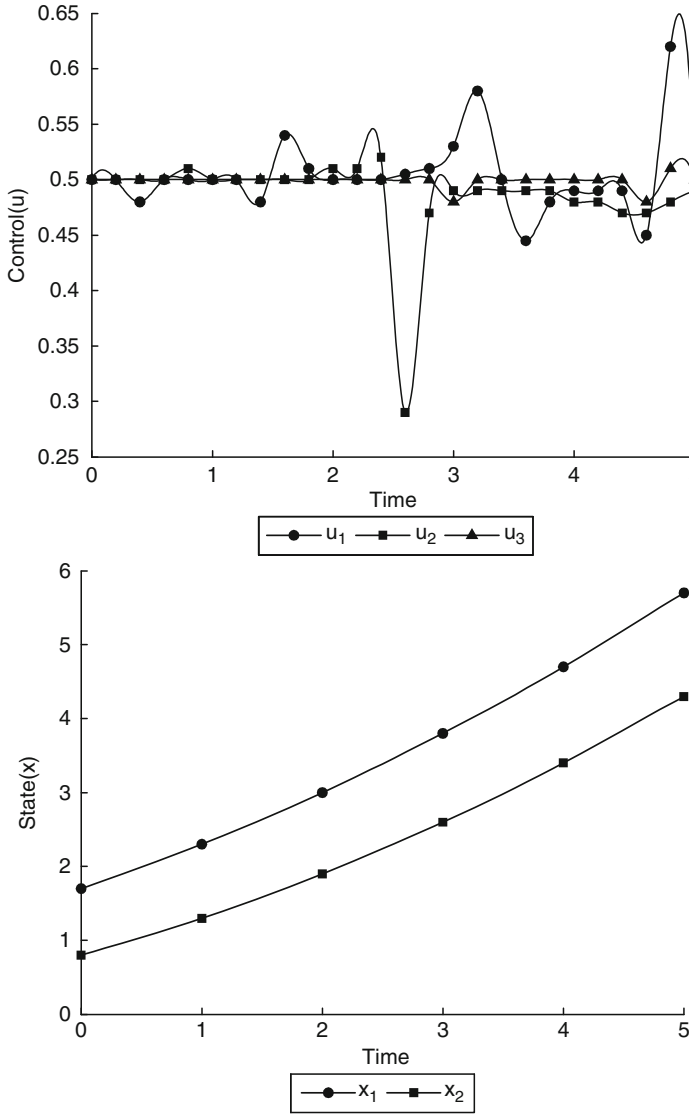


Fig. 6.6 u^* vs. time and x^* vs. time without time shifts

Figure 6.6 shows the corresponding numerical solution of *DVI(F, f, Ψ, U, x_0, t_0, t_f, $\tau = 0$)*. The fixed-point-descent-in-Hilbert-space algorithm converged in 12 fixed-point iterations, and each of the subproblems converged in nine or fewer iterations.

6.7 Exercises

1. By analogy to the finite-dimensional case, define the notion of a differential quasivariational inequality
2. Explain why a generalized differential Nash equilibrium is equivalent to an appropriately defined differential quasivariational inequality.
3. Establish that, for appropriate regularity conditions, $x(u, u_\tau, t)$ exists.
4. Establish that, for appropriate regularity conditions, $x(u, u_\tau, t)$ is unique.
5. Consider the following *DVI(F, f, Ψ, U, x_0, t_0, t_f)* involving two controls and two states:

$$u \in \left(L^2[0,1]\right)^2$$

$$x \in \left(\mathcal{H}^1[0,1]\right)^2$$

$$x(t_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\left[t_0, t_f\right] = [0,5]$$

$$F(x,u) = \begin{pmatrix} F_1(x,u) \\ F_2(x,u) \end{pmatrix} = \begin{pmatrix} (x_1)^2 - u_1 \\ x_2 - u_2 \end{pmatrix}$$

$$f(x,u) = \begin{pmatrix} f_1(x,u) \\ f_2(x,u) \end{pmatrix} = \begin{pmatrix} x_1 - u_1 + (u_2)^2 \\ x_2 + (u_1)^2 - u_2 \end{pmatrix}$$

$$U = \{u : 0 \le u_1 \le 1;\ 0 \le u_2 \le 1;\ 0 \le u_3 \le 1\}$$

   Solve this problem using a fixed-point algorithm in continuous time.
6. Repeat Exercise 5 using a gap function in continuous time.
7. Repeat Exercise 5 using a discrete-time approximation.
8. Consider the following *DVI(F, f, Ψ, U, x_0, t_0, t_f, τ)* involving two controls and two states:

$$u \in \left(L^2[0,1]\right)^2$$

$$x \in \left(\mathcal{H}^1[0,1]\right)^2$$

$$x(t_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\left[t_0, t_f\right] = [0,5]$$

$$F(x,u,u_\tau) = \begin{pmatrix} F_1(x,u) \\ F_2(x,u) \end{pmatrix} = \begin{pmatrix} (x_1)^2 - u_1 + u_2(t-4) \\ (x_2)^2 + (u_1)^2 - u_2 \end{pmatrix}$$

$$f\ (x,u) = \begin{pmatrix} f_1(x,u) \\ f_2(x,u) \end{pmatrix} = \begin{pmatrix} -u_1 \\ -u_2 \end{pmatrix}$$

$$U\ = \{u : 0 \leq u_1 \leq 1;\ 0 \leq u_2 \leq 1;\ 0 \leq u_3 \leq 1\}$$

where $u_2\ (t-4)$ denotes a control variable with a fixed time shift. Solve this problem using a fixed-point algorithm in continuous time.

List of References Cited and Additional Reading

Adams, R. A. (1970). Equivalent norms for Sobolev spaces. Proceedings of the American Mathematical Society 24(1), 63–66.

Aubin, J. P. and A. Cellina (1984). *Differential Inclusions*. New York: Springer-Verlag.

Bernstein, D., T. L. Friesz, R. L. Tobin, and B. W. Wie (1993). A variational control formulation of the simultaneous route and departure-time equilibrium problem. *Proceedings of the International Symposium on Transportation and Traffic Theory*, 107–126.

Bressan, A. and B. Piccoli (2007). *An Introduction to the Mathematical Theory of Control*. Springfield, MO: American Institute of Mathematical Sciences.

Brouwer, L. E. J. (1910). Ueber eineindeutige, stetige transformationen von flächen in sich. *Mathematische Annalen 69*(2), 176–180.

Browder, F. E. (1968). The fixed point theory of multivalued mappings in topological vector spaces. *Mathematische Annalen 177*, 283–301.

Friesz, T., D. Bernstein, and R. Stough (1996). Dynamic systems, variational inequalities, and control-theoretic models for predicting time-varying urban network flows. *Transportation Science 30*(1), 14–31.

Friesz, T. and R. Mookherjee (2006). Differential variational inequalities with state-dependent time shifts. *Transportation Research Part B*.

Friesz, T. L., D. Bernstein, T. Smith, R. Tobin, and B. Wie (1993). A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research 41*(1), 80–91.

Fukushima, M. (1992). Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming 53*, 99–110.

Isaacs, R. (1965). *Differential Games*. New York: Dover.

Kachani, S. and G. Perakis (2002a). *A fluid model of dynamic pricing and inventory management for make to stock manufacturing systems*. Technical report, Sloan School of Management, MIT.

Kachani, S. and G. Perakis (2002b). *Fluid dynamics models and their application in transportation and pricing*. Technical report, Sloan School of Management, MIT.

Konnov, I. V. and S. Kum (2001). Descent methods for mixed variational inequalities in a Hilbert space. *Nonlinear Analysis: Theory, Methods and Applications 47*(1), 561–572.

Konnov, I. V., S. Kum, and G. M. Lee (2002). On convergence of descent methods for variational inequalities in a Hilbert space. *Mathematical Methods of Operations Research 55*, 371–382.

Kwon, C., T. Friesz, R. Mookherjee, T. Yao, and B. Feng (2009). Non-cooperative competition among revenue maximizing service providers with demand learning. *European Journal of Operational Research 197*(3), 981–996.

Minoux, M. (1986). *Mathematical Programming: Theory and Algorithms*. New York: John Wiley.

Mookherjee, R. and T. Friesz (2008). Pricing, allocation, and overbooking in dynamic service network competition when demand is uncertain. *Production and Operations Management 14*(4), 1–20.

Patriksson, M. (1997). Merit functions and descent algorithms for a class of variational inequality problems. *Optimization 41*, 37–55.

Peng, J.-M. (1997). Equivalence of variational inequality problems to unconstrained minimization. *Mathematical Programming 78*, 347–355.

Perakis, G. (2000). *The dynamic user equilibrium problem through hydrodynamic theory*. Sloan School of Management, MIT, preprint.

Pshenichnyi, B. N. (1971). *Necessary Conditions for an Extremum*. New York: Marcel Dekker.

Todd, M. J. (1976). *The Computation of Fixed Points and Applications*. New York: Springer-Verlag.

Yamashita, N., K. Taji, and M. Fukushima (1997). Unconstrained optimization reformulations of variational inequality problems. *Journal of Optimization Theory and Applications 92*(3), 439–456.

Zhu, D. L. and P. Marcotte (1994). An extended descent framework for variational inequalities. *Journal of Optimization Theory and Applications 80*(2), 349–366.

Zhu, D. L. and P. Marcotte (1998). Convergence properties of feasible descent methods for solving variational inequalities in Banach spaces. *Computational Optimization and Applications 10*(1), 35–49.

Chapter 7
Optimal Economic Growth

The theory of optimal economic growth is a branch of economic theory that makes direct and sophisticated use of the theory of optimal control. As such, the models of optimal economic growth that have been devised and reported in the economics literature are relatively easy for a person who has mastered the material of Chapters 3 and 4 of this book to comprehend. Among other things, this chapter shows how aspatial optimal economic growth theory may be extended to study optimal growth of interdependent regions in a national economy. Moreover, working through the analyses presented in this chapter provides a means of assessing and improving one's mastery of the key mathematical concepts from the theory of optimal control that were introduced in previous chapters, especially the analysis and interpretation of optimality conditions and singular controls.

Multiregional optimal growth models are generally notationally complicated, and those presented in this chapter are no exception. For that reason, the first multiregional model of optimal econmomic growth considered here is based on a very simple model of production that is quite tractable even though it is theoretically somewhat naive and outdated. In particular, Section 7.2 presents an extremely detailed analysis of that model's optimality conditions, study of which will allow the reader to become familiar with the style and depth of analysis that must be conducted when a new optimal control model is created. From there we go on to discuss, in Section 7.3, a more advanced model of optimal regional growth, similar to the growth theory application introduced in Chapter 1, that will provide the professional economist with some practical knowledge about how to numerically solve models of optimal economic growth using a discrete time/mathematical programming approach.

The following is a preview of the principal topics covered in this chapter:

Section 7.1: Alternative Models of Optimal Economic Growth. In this section, we present Ramsey's famous model of optimal growth. We also formulate a model of optimal economic growth based on the Harrod-Domar growth dynamics. Additionally, we present the optimal control problem at the heart of the neoclassical theory of optimal growth.

Section 7.2: Optimal Regional Growth Based on the Harrod-Domar Model. In this section, we show how the Harrod-Domar model may be disaggregated to create a multiregion optimal growth model.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_7, © Springer Science+Business Media, LLC 2010

Section 7.3: A Computable Theory of Regional Public Investment Allocation. In this section, we show how public investment may be optimally allocated among regions without invoking any assumptions regarding economies of scale.

7.1 Alternative Models of Optimal Economic Growth

There are alternative theories of optimal economic growth. We now review three of these that take the form of optimal control problems.

7.1.1 Ramsey's 1928 Model

As we noted in Chapter 1, Ramsey (1928) proposed the idea of a *bliss point*, an accumulation point of a sequence of consumption decisions representing the nonattainable, ideal consumption goal of the consumer. The bliss point, $B > 0$, has the same units as utility and obeys

$$B = \sup [U(c) : c \geq 0]$$

where c is the consumption level of a representative member of society and $U(c)$ is the utility experienced as a result of that consumption. Because there are N identical members of society, maximization of social welfare is assured by

$$\max J = \int_0^\infty [U(c) - B]\, dt \tag{7.1}$$

The relevant dynamics are obtained from a neoclassical production function

$$Y = F(K, L) \tag{7.2}$$

where Y is output and K and L are capital and labor inputs, respectively. The neoclassical nature of (7.2) means that $F(.,.)$ is homogeneous of degree one; that is

$$F(\alpha K, \alpha L) = \alpha F(K, L)$$

for α a positive scalar. Taking labor (population) to be fixed, per capita output may be expressed as

$$\begin{aligned} y &= \frac{Y}{L} = \frac{1}{L} F(K, L) \\ &= F\left(\frac{K}{L}, 1\right) = f(k) \end{aligned}$$

where

$$k \equiv \frac{K}{L} \quad \text{and} \quad f(k) \equiv F\left(\frac{K}{L}, 1\right)$$

Thus, we have

$$\frac{dk}{dt} = \frac{d}{dt}\left(\frac{K}{L}\right) = \frac{1}{L}\frac{dK}{dt} = \frac{I}{L}$$

from which we obtain

$$\begin{aligned} \frac{dk}{dt} &= \frac{F(K, L)}{L} - \frac{C}{L} - \frac{\delta K}{L} \\ &= f(k) - c - \delta k \end{aligned} \tag{7.3}$$

where

$$I = Y - C - \delta K$$

is investment, c is the per capita consumption rate, $(1 - c)$ is the savings rate, C is total consumption, and δ is the rate of depreciation of capital. Obviously, k is per capita capital. Note that per capita output $f(k)$ is expressed in terms of the single state variable k, a fact made possible by the homogeneous-of-degree-one property of the production function. In deriving (7.3), we make use of well-known macroeconomic identities relating the rate of change of capital stocks dK/dt, investment I, consumption C, and capital depreciation δK:

$$\begin{aligned} \frac{dK}{dt} &= I - \delta K \\ &= Y - C - \delta K \end{aligned}$$

The above development allows us to state Ramsey's optimal growth model in the following form:

$$\max J = \int_0^\infty [U(c) - B]\, dt \tag{7.4}$$

subject to

$$\frac{dk}{dt} = f(k) - \delta k - c \tag{7.5}$$

$$k(0) = k_0 \tag{7.6}$$

The model (7.4), (7.5), and (7.6) is an optimal control problem with state variable k and control variable c.

7.1.2 Optimal Growth with the Harrod-Domar Model

Ramsey's work can be criticized from the points of view that population is not time varying that the concept of a bliss point contradicts the nonsatiation axiom of

utility theory. We may quite easily overcome these weaknesses, and we now do so for an especially simple production technology. Specifically, we imagine a production function of the form

$$Y(t) = \sigma K(t) \tag{7.7}$$

where $Y(\cdot)$ is the aggregate output, σ is the output-capital ratio, $K(\cdot)$ is aggregate capital, and t is a continuous-time variable. Relationship (7.16) is the basis of the well-known Harrod-Domar model, for which the output-capital ratio is argued to be constant on a so-called balanced growth path [see Hahn and Matthews (1964)]. Thus, one may write

$$\frac{dY(t)}{dt} = \sigma \frac{dK(t)}{dt} = \sigma I(t) \tag{7.8}$$

where $I(\cdot)$ denotes total investment. We also know that

$$\begin{aligned} \frac{dK}{dt} &= I - \delta K \\ &= Y - C - \delta K \end{aligned}$$

where I is investment and C is total consumption. If we again assume labor to be fixed, we have

$$\frac{dk}{dt} = \sigma k - c - \delta k$$

where c is per capita consumption, δ is the rate of depreciation of capital and k is per capita capital. Upon introducing a constrant nominal interest rate ρ, we give the following as a model of optimal economic growth:

$$\max J = \int_0^\infty \exp(-\rho t) U(c)\, dt \tag{7.9}$$

subject to

$$\frac{dk}{dt} = (\sigma - \delta) k - c \tag{7.10}$$

$$k(0) = k_0 \tag{7.11}$$

The model (7.9), (7.10), and (7.11) is an optimal control problem with state variable k and control variable c.

7.1.3 Neoclassical Optimal Growth

The weaknesses of Ramsey's model may also be overcome for a more general class of production functions involving capital and labor inputs and constant scale

economies of production; the resulting model is known as the neoclassical optimal growth model. One of the clearest and most succinct expositions of the neoclassical theory of optimal economic growth is contained in the book by Arrow and Kurz (1970). To express the neoclassical model of optimal growth, as explained in Chapter 1, we postulate that population has a constant proportionate growth rate π obeying

$$\frac{1}{L}\frac{dL}{dt} = \pi \Longrightarrow L = L_0 \exp(-\pi t) \tag{7.12}$$

where $L_0 = L(0)$ is the initial population (labor force). We then construct dynamics for per capita capital $k = K/L$ in a fashion highly similar to that employed for Ramsey's model. In particular we write

$$\begin{aligned}
\frac{dk}{dt} &= \frac{d}{dt}\left(\frac{K}{L}\right) = \frac{1}{L}\frac{dK}{dt} - \frac{K}{L}\frac{1}{L}\frac{dL}{dt} \\
&= \frac{1}{L}I - \pi k \\
&= \frac{(1-c)Y - \delta K}{L} - \pi\frac{K}{L} \\
&= \frac{F(K,L)}{L} - \frac{C}{L} - \delta\frac{K}{L} - \pi\frac{K}{L} \\
&= f(k) - c - \delta k \\
&= f(k) - c - \delta k - \pi k
\end{aligned}$$

where

$$f(k) = F\left(\frac{K}{L}, 1\right)$$

as before. Consequently, the neoclassical optimal growth model is

$$\max J = \int_0^\infty \exp(-\rho t)\, U(c)\, dt \tag{7.13}$$

subject to

$$\frac{dk}{dt} = f(k) - (\delta + \pi)k - c \tag{7.14}$$

$$k(0) = k_0 \tag{7.15}$$

where ρ is the constant nominal interest rate. The model (7.13), (7.14), and (7.15) is again an optimal control problem. The neoclassical optimal growth model is itself open to criticism, especially as regards its assumption of a constant returns to scale technology.

7.2 Optimal Regional Growth Based on the Harrod-Domar Model

In this section, we consider two especially easy-to-analyze models of optimal economic growth that are interesting in their own right and that provide an introduction to the type of thinking characteristic of dynamic continuous-time optimization modeling applications. We assume an economy with a private sector and a public sector comprising n different geographical regions. For each sector within each region, the production function is assumed to be of the Harrod-Domar form introduced above:

$$Y(t) = \sigma K(t) \tag{7.16}$$

where $Y(\cdot)$ is the aggregate output, σ is the output-capital ratio, $K(\cdot)$ is aggregate capital, and t is a continuous-time variable. Because the output-capital ratio is constant on the balanced growth path, we have

$$\frac{dY(t)}{dt} = \sigma \frac{dK(t)}{dt} = \sigma I(t) \tag{7.17}$$

where $I(\cdot)$ denotes total investment.

To construct an n-region model let us introduce the following notation:

$$\begin{aligned}
Y_i &= \text{ total income of region } i \\
\sigma_i &= \text{output-capital ratio for private investment in region } i \\
\delta_i &= \text{ output-capital ratio for public investment in region } i \\
r &= \text{uniform income tax rate for all regions controlled} \\
&\quad \text{by the central government} \\
s_i &= \text{savings ratio in region } i \\
\psi_{ij} &= \text{fraction of private savings generated in region } j \\
&\quad \text{that is transferred to region } i \\
a_{ij} &= \text{fraction of transferred private savings } \textit{not} \text{ consumed} \\
&\quad \text{in the transportation process} \\
\mu_{ij} &= \text{fraction of public savings generated in region } j \text{ that} \\
&\quad \text{is transferred to region } i \\
b_{ij} &= \text{fraction of transferred public savings funds not consumed} \\
&\quad \text{in the transportation process}
\end{aligned}$$

These definitions allow dynamical descriptions for the evolution of regional incomes with respect to time to be written down. One point of view is to assume that from the jth region's income, Y_i, a fraction $(1 - s_j)$ is the consumption allowance of region j, including both public and private consumption. The remainder $s_j Y_j$ is taxed at the rate $r(t)$, leaving $(1 - r(t))s_j Y_j$ to be privately controlled and $r(t)s_j Y_j$

to be publicly controlled. Some fraction of these region j funds are transferred to region i. In fact $(1 - r(t))\psi_{ij} a_{ij} s_j Y_j$ describes the amount of funds transferred to region i from the jth region's public sector. It clearly follows that the rate of change of the ith region's income with respect to time can be formulated as

$$\frac{dY_i(t)}{dt} = (1 - r(t))\sigma_i \sum_{j=1}^{n} \psi_{ij} a_{ij} s_j Y_j + r(t)\delta_i \sum_{j=1}^{n} \mu_{ij}(t) b_{ij} s_j Y_j \qquad i \in [1, n] \tag{7.18}$$

Expression (7.18) of course gives rise to a set of simultaneous differential equations, since (7.18) holds for every region $i \in [1, n]$. Moreover, expression (7.18) can be considered a generalization of the dynamics proposed by Sakashita (1967a).

A differential point of view has been proposed by Friesz and Luque (1987). They assume that the income of region j is divided, by means of the tax rate $r(t)$, between the public and private sectors: the after-tax income for the private sector is $(1 - r(t))Y_j(t)$, and the corresponding public sector income is $r(t)Y_j(t)$. The relevant private and public sector investment funds are generated through a private savings ratio, s_j, and a public saving ratio, v_j. As before, these funds are allocated and transferred to the private and public investment pools of other regions, finally obtaining

$$\frac{dY_i(t)}{dt} = (1 - r(t))\sigma_i \sum_{j=1}^{N} \psi_{ij} a_{ij} s_j Y_j + r(t)\delta_i \sum_{j=1}^{N} \mu_{ij}(t) b_{ij} v_j Y_j \qquad i \in [1, n] \tag{7.19}$$

The initial conditions for both sets of dynamics (7.18) and (7.19) are $Y_i(0) = Y_i^0$, known constants for each region $i \in [1, n]$.

The income tax rate is such that

$$0 \leq r(t) \leq \theta \tag{7.20}$$

where θ is a known constant and $0 < \theta < 1$. Note also the transfer allocation controls μ_{ij}, for all time $t \in [t_0, t_f]$, satisfy

$$\sum_{i=1}^{n} \mu_{ij}(t) = 1 \qquad i \in [1, n] \tag{7.21}$$

One does not have to worry about circularities in public sector transfers because the b_{ij} are positive, but less than one; consequently, transfers have implicit nonzero costs, $1 - b_{ij}$, which ensure that the optimal solution is well behaved. Circularities with respect to transfers in the private sector can exist in principle, since that sector is composed of independent decision makers, each of whom makes claims upon only a part of the income of the private sector and each of whom possesses distinct

views regarding the best investment opportunities. Nonetheless, we treat ψ_{ij} as an exogenously supplied parameter and require that

$$\sum_{i=1}^{n} \psi_{ij} = 1 \qquad i \in [1, n] \tag{7.22}$$

so that the aforementioned issue of private sector transfer circularities is not a problem in our analysis.

A very general objective function that is a scalarized version of a vector objective function is

$$J = w_1 \sum_{i=1}^{n} \alpha_i Y_i(t_f) + w_2 \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} \beta_i Y_i(t) dt \tag{7.23}$$

The intent is, of course, to maximize J. The weights α_i and β_i describe different relative values for their respective scalar objectives: the ith region's end of period income $Y_i(t_f)$ and the ith region's income time stream $\int_{t_0}^{t_f} \exp(-\rho t) Y_i(t) dt$. The weights w_1 and w_2 determine the relative importance of aggregate terminal versus aggregate time stream benefits. Thus (7.23) is really based on a scalarization of the following vector objective function

$$Z = [Y_1(t_f), \ldots, Y_n(t_f); \int_{t_0}^{t_f} \exp(-\rho t) Y_1(t) dt, \ldots, \int_{t_0}^{t_f} \exp(-\rho t) Y_n(t) dt] \tag{7.24}$$

As in Section 7.1, the entity ρ is a constant nominal interest rate used to calculate present values.

We now show that maximization of (7.23) subject to (7.19), (7.20), and (7.21) includes previous models reported in the literature as special cases. The first paper on the subject of optimal regional public investment allocation from a mathematical perspective appears to be Rahman (1963a) using a discrete-time framework. Intriligator (1964) set forth a continuous-time formulation of Rahman's approach, but made a mistake in its solution as Rahman (1963a) noted. Eventually, the original Rahman model was thoroughly solved by Takayama (1967). The Rahman model may be obtained from our formulation by making these simplifications:

1. Assume there is only one sector, namely the public sector, since interregional transfers of funds are centrally planned. This implies that $r(t) = 1$ for all $t \in [t_0, t_f]$.
2. Also assume zero transportation costs; thus $b_{ij} = 1$ for all i and j.
3. Further assume that public savings are first consolidated and then distributed, which is equivalent to requiring $\mu_{ij}(t) = \mu_i(t)$ for all j.

As a consequence the relevant constrained dynamics, for each $i \in [1, n]$ and all $t \in [t_0, t_f]$, are

$$\frac{dY_i(t)}{dt} = \delta_i \mu_i(t) \sum_j v_j Y_j(t) = \delta_i \mu_i(t) \sum_j v_j Y_j(t) \qquad i \in [1, n] \tag{7.25}$$

$$Y_i(t_0) = Y_i^0 \qquad i \in [1, n] \tag{7.26}$$

$$\sum_i \mu_i(t) = 1 \tag{7.27}$$

$$\mu_i(t) \geq 0 \qquad i \in [1, n] \tag{7.28}$$

The objective of Rahman's model is to maximize national income at the terminal time $t = t_f$; that objective function is an obvious special case of (7.23), and we write it as

$$J = \sum_{i=1}^{n} Y_i(t_f) \tag{7.29}$$

It is important to observe that Intriligator (1964) also analyzed the alternative objective of maximizing undiscounted per capita consumption for which he assumes exponential population growth at the proportionate rate m; thus for Intriligator

$$J = \int_{t_0}^{t_f} \exp(-mt) \sum_{i=1}^{n} (1 - v_i) Y_i(t) dt \tag{7.30}$$

Expression (7.30) is clearly a special case of our general objective function (7.23). In his analysis of the problem, Takayama (1967) introduced a constant nominal interest rate ρ; under the assumption $\rho + m > 0$, he considers the objective

$$\max J = \int_{t_0}^{t_f} \exp[-(m + \rho)t] \sum_{i=1}^{n} (1 - v_i) Y_i(t) dt \tag{7.31}$$

which is obviously another special case of (7.23).

Transportation is introduced for the first time by Datta-Chaudhuri (1967). He assumes an *ex ante* neoclassical production function, without techn0logical change and with the usual convexity (diminishing returns) assumption; therefore

$$Y_i = F_i(K_i, L_i) = L_i \cdot F_i\left(\frac{K_i}{L_i}, 1\right) = L_i \cdot f_i\left(\frac{K_i}{L_i}\right) \tag{7.32}$$

for each region $i \in [1, n]$. One of the cases analyzed by Datta-Chaudhuri (1967), the one of unlimited supply of labor, which would arise for example in the first stages of development of an economy, can be easily fitted into our framework. Since labor is unlimited, K_i/L_i will be maintained at the constant optimal level k_i; that is, $L_i = K_i/k_i$ and

$$Y_i = \frac{K_i f_i(k_i)}{k_i} = \delta_i K_i \qquad \delta_i = \frac{f_i(k_i)}{k_i} \tag{7.33}$$

Savings are assumed to be a fraction of profits after labor wages have been paid out. If V_i denotes total public savings

$$\begin{aligned} V_i(t) &= v_i\,(Y_i - w_i L_i) \\ &= v_i L_i\,(f_i(k_i) - w_i) \\ &= \left[v_i\left(\frac{f_i(k_i) - w_i}{k_i}\right)\right] K_i \\ &= \left[v_i\left(\frac{f_i(k_i) - w_i}{k_i \delta_i}\right)\right] Y_i \end{aligned} \tag{7.34}$$

for each region $i \in [1, n]$. By redefining v_i as

$$v_i\left(\frac{f_i(k_i) - w_i}{k_i \delta_i}\right) \tag{7.35}$$

we get

$$\frac{dY_i}{dt} = \delta_i \sum_{j=1}^{n} \mu_{ij}(t) b_{ij} v_j Y_j(t) \qquad i \in [1, n] \tag{7.36}$$

In his paper, Datta-Chandhuri analyzes a two-region model, and makes the following assumptions:

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1-b \\ 1-b & 1 \end{pmatrix}$$

$$\begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{pmatrix} = \begin{pmatrix} 1-\mu_1 & \mu_1 \\ \mu_2 & 1-\mu_2 \end{pmatrix}$$

Thus, for the two-region case, expression (7.36) specializes to Datta-Chaudhuri's dynamic equations:

$$\begin{aligned} \frac{dY_1}{dt} &= \delta_1[(1-\mu_1(t))v_1 Y_1 + (1-b)\mu_2 v_2 Y_2] \\ \frac{dY_2}{dt} &= \delta_2[(1-b)\mu_1(t) v_1 Y_1 + (1-\mu_2) v_2 Y_2] \end{aligned}$$

The constraints on the controls are

$$\mu_1(t), \mu_2(t) \in [0, 1] \qquad t \in \left[t_0, t_f\right] \tag{7.37}$$

The objective considered by Datta-Chaudhuri (1967) is to minimize the time to reach a certain level of total capital stock; that is

$$\min J = \int_{t_0}^{t_f} dt \quad \text{subject to} \quad \sum_{i=1}^{n} \frac{Y_i(t_f)}{\delta_i} \geq \overline{K} \tag{7.38}$$

A similar objective is that of maximizing total capital stock at the end of the planning period:

$$\max J = \sum_{i=1}^{n} \frac{Y_i(t_f)}{\delta_i} \tag{7.39}$$

This objective is clearly a special case of the general objective function (7.23) and so, in light of the preceding development, we may consider Datta-Chandhuri's model a special case of our general model.

Domazlicky (1977) uses the same two-region model introduced above, including the same production function $Y(t) = \sigma K(t)$, and considers the objective function proposed by Rahman (1963a), namely

$$\max J = \sum_{i=1}^{n} Y_i(t_f) \tag{7.40}$$

Consequently, the Domazlicky model is another special case of our general model. Domazlicky argues that Rahman's and Datta-Chandhuri's results are not comparable because the latter uses a neoclassical production function. However, we have seen that under the assumption of unlimited supply of labor, the neoclassical production function may be re-expressed in Harrod-Domar form. Differences arise, in any case, from the authors' consideration of different objective functions: minimization of the time in which a certain level of total capital stock is reached versus maximization of the end of period sum of regional incomes. Clearly, the objective function (7.23) of our general model will allow both of these objectives to be considered simultaneously if desired.

Two sectors, public vs. private, appear for the first time in Sakashita (1967a), who considers two distinct types of public investment. The first type, social overhead investment, has indirect effects on changes in the productivity of already installed capital and on the allocation of private funds to the different regions. The second type, productive government expenditure, has direct effects on both the levels of capital stock and on income. We will limit our attention to productive government expenditure models; this type of model has been extended to n regions by Ohtsuki (1971), and it will be his version that we analyze and to which will draw contrasts. For Ohtsuki (1971), transportation costs are assumed to be zero and private and public funds are first consolidated and then distributed among regions; that is, for all $i \in [1, n]$ and $j \in [1, n]$:

$$\begin{aligned} b_{ij} &= 1 \\ \psi_{ij} &= \psi_i \\ \mu_{ij} &= \mu_i \end{aligned}$$

The regional income dynamics are

$$\frac{dY_i}{dt} = (1 - r(t))\psi_i \sigma_i \sum_{j=1}^{n} s_j Y_j + r(t)\mu_i(t)\delta_i \sum_{j=1}^{n} s_j Y_j \qquad i \in [1, n] \tag{7.41}$$

Ohtsuki generalizes the unweighted end-of-period income

$$\max J = \sum_{i=1}^{n} Y_i(t_f) \tag{7.42}$$

to

$$\max J = \sum_{i=1}^{n} \alpha_i Y_i(t_f) \tag{7.43}$$

In the two-region case the dynamical equations are

$$\frac{dY_1}{dt} = (1 - r(t))\psi\sigma_1(s_1Y_1 + s_2Y_2) + r(t)\mu(t)\delta_1(s_1Y_1 + s_2Y_2) \tag{7.44}$$

$$\frac{dY_2}{dt} = (1 - r(t))(1 - \psi)\sigma_2(s_1Y_1 + s_2Y_2) + r(t)(1 - \mu(t))\delta_2(s_1Y_1 + s_2Y_2) \tag{7.45}$$

The following are

$$0 \le r(t) \le \theta \quad 0 \le \mu(t) \le 1 \quad \forall t \in [t_0, t_f] \tag{7.46}$$

control constraints.

Proceeding in a similar way, Friesz and Luque (1987) generalize the Sakashita model to take into account different savings ratio in the public and private sectors; as an objective, they use maximization of the present value of total income:

$$\max J = \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} Y_i(t)dt \tag{7.47}$$

where the nominal interest rate is ρ. The corresponding equations of motion analyzed by Friesz and Luque (1987) are

$$\frac{dY_i(t)}{dt} = (1 - r(t))\psi_i\sigma_i \sum_{j=1}^{n} s_jY_j + r(t)\mu_i(t)\delta_i \sum_{j=1}^{n} v_jY_j \qquad i \in [1, n]$$

For two regions, the dynamical equations studied by Friesz and Luque (1987) are

$$\frac{dY_1(t)}{dt} = (1 - r(t))\psi\sigma_1(s_1Y_1 + s_2Y_2) + r(t)\mu(t)\delta_1(v_1Y_1 + v_2Y_2) \tag{7.48}$$

$$\frac{dY_2(t)}{dt} = (1 - r(t))(1 - \psi)\sigma_2(s_1Y_1 + s_2Y_2) + r(t)(1 - \mu(t))\delta_2(v_1Y_1 + v_2Y_2)$$

By virtue of the above presentation, the Ohtsuki model given by (7.43), (7.44), and (7.46) and the Sakashita-Friesz model given by (7.46), (7.47), and (7.48) are both clearly special cases of the general model given by (7.20), (7.21), (7.22), and (7.23).

7.2.1 Tax Rate as the Control

The general model introduced above and its variants, which we have described in the previous section, are expressed as optimal control problems; their solution requires the application of the Pontryagin maximum principle presented in previous chapters. To appreciate fully the nature of solutions obtained from application of this type of necessary condition to models of the type described in the previous section, it is instructive to first consider a somewhat simplified single objective problem which is a specified case of the general model. In particular, let the only control variable be the tax rate and take the public investment share μ_i to be a constant for all regions $i \in [1, n]$. The model, henceforth called problem $P1$ for short, is the following:

$$\max J = \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} \beta_i Y_i(t) dt \tag{7.49}$$

subject to

$$\frac{dY_i}{dt} = \sigma_i \psi_i (1 - r(t)) \sum_{j=1}^{n} s_j Y_j + \delta_i \mu_i r(t) \sum_{j=1}^{n} v_j Y_j \qquad i \in [1, n] \tag{7.50}$$

$$Y_i(t_0) = Y_i^0 > 0 \qquad i \in [1, n] \tag{7.51}$$

$$0 \leq r(t) \leq \theta \leq 1 \tag{7.52}$$

Here all notation is as previously defined. The β_i are weights attached to individual regional incomes so that $\sum_i \beta_i Y_i$ is a scalarized version of the vector function

$$[Y_1(t), \ldots, Y_n(t)] \tag{7.53}$$

Of course, in (7.50) we have the implicit fact that

$$\sum_{i=1}^{n} \mu_i = \sum_{i=1}^{n} \psi_i = 1$$

By introducing the matrices

$$A = \left(\sigma_i \psi_i s_j\right) \quad i \in [1, n] \quad j \in [1, n]$$
$$B = \left(\delta_i \mu_i v_j\right) \quad i \in [1, n] \quad j \in [1, n]$$

and the vectors

$$Y = (Y_i, \ldots, Y_n)^T \quad \beta = (\beta_1, \ldots, \beta_n)^T \tag{7.54}$$

we can place the specialized model described by (7.49), (7.50), (7.51), and (7.52) into the following form:

$$\max J = \int_{t_0}^{t_f} \exp(-\rho t) \beta^T Y(t) dt \tag{7.55}$$

subject to

$$\frac{dY}{dt} = [A + r(t)(B - A)]Y \tag{7.56}$$

$$Y(t_0) = Y^0 \in \Re^n_{++} \tag{7.57}$$

$$0 \leq r(t) \leq \theta \leq 1 \tag{7.58}$$

$$0 \leq r(t) \leq \theta \leq 1 \tag{7.59}$$

where $Y^0 = (Y_1^0, \ldots, Y_n^0)^T$ is a given positive vector. The Hamiltonian for the problem described by (7.55), (7.56), (7.57), (7.58), and (7.59) is

$$H = \exp(-\rho t)\beta^T Y + \lambda\,[A + r(B - A)]\,Y \tag{7.60}$$

where λ denotes a vector of adjoint variables. The adjoint equations and transversality conditions are:

$$\frac{d\lambda}{dt} = -H_Y = -\exp(-\rho t)\beta^T - \lambda\,[A + r(B - A)] \tag{7.61}$$

$$\lambda(t_f) = 0 \quad H(t) \text{ free} \tag{7.62}$$

In order to apply the Pontryagin maximum principle we note that

$$H_r = \lambda(B - A)Y \tag{7.63}$$

Consequently, the extremal tax rate is given by

$$r^* = \begin{cases} \theta \text{ if } \lambda(B - A)Y > 0 \\ 0 \text{ if } \lambda(B - A)Y < 0 \\ r_s \text{ if } \lambda(B - A)Y + 0 \end{cases} \tag{7.64}$$

The case $\lambda(B - A)Y = 0$, if it occurs for an arc of time, corresponds to a singular control r_s.

The value of the control on a singular arc is found by taking successive time derivatives of the gradient of the Hamiltonian H_r and requiring that these derivatives vanish. For this particular case the vanishing of H_r yields the following expression for the singular control:

$$r^*(t) = -\frac{\{\lambda[[B, A]^*, A]^* + \exp(-\rho t)\beta^T[(B - A)(\rho I - A) - [B, A]^*]\}Y}{\{\lambda[[B, A]^*, (B - A)]^* - \exp(-\rho t)\beta^T(B - A)(B - A)\}Y} \tag{7.65}$$

where $[C, D]^*$ represents the commutation operator $CD - DC$ for any pair of square matrices C and D. In order that the singular control (7.65) maximize the Hamiltonian as required by the Pontryagin conditions, we must invoke the generalized convexity condition for singular arcs:

$$-\frac{\partial}{\partial r}\left[\frac{d^2 H_r}{dt^2}\right] \leq 0 \tag{7.66}$$

For our case (7.66) becomes

$$\{\lambda[[B, A]^*, (B - A)]^* - \exp(-\rho t)\beta^T (B - A)(B - A)\}Y \geq 0 \tag{7.67}$$

The complete optimal policy for this scalar example will consist of a combination of so-called "bang-bang" controls and singular controls described by (7.64) and (7.65). In fact, more general vector problems that consider terminal as well as time stream benefits and public investment shares as well as the tax rate as control variables may be formulated. These too, as we will see, may possess both singular and bang-bang controls.

The interpretation of (7.64) is straightforward. We see that the different possibilities are

$$pBY \gtreqless pAY \tag{7.68}$$

or equivalently

$$\left(\sum_{i=1}^{n} \lambda_i \delta_i \mu_i\right)\left(\sum_{j=1}^{n} v_j Y_j\right) - \left(\sum_{i=1}^{n} \lambda_i \sigma_i \psi_i\right)\left(\sum_{j=1}^{n} s_j Y_j\right) \gtreqless 0 \tag{7.69}$$

Per unit of tax rate, $\sum_j s_j Y_j$ and $\sum_j v_j Y_j$ are the total amounts of investment funds lost by the private sector and gained by the public sector, respectively. Those losses or gains are allocated to different regions according to the parameters ψ_i and μ_i, $i = 1, \ldots, n$, where they produce changes in income given by

$$\Delta Y_i^p = \sigma_i \psi_i \left(\sum_{j=1}^{n} s_j Y_j\right) \quad i \in [1, n] \tag{7.70}$$

$$\Delta Y_i^g = \delta_i \mu_i \left(\sum_{j=1}^{n} v_j Y_j\right) \quad i \in [1, n] \tag{7.71}$$

In the private sector (p) these changes are income losses; in the public sector (g) they are income gains. These income losses and gains should not be compared directly; instead, the appropriate comparison is in terms of their effects on the objective function. Along an optimal path, it is well known that the adjoint variables satisfy

$$\lambda_i = \frac{\partial J}{\partial Y_i} \quad i \in [1, n] \tag{7.72}$$

Therefore

$$\lambda_i \sigma_i \psi_i \left(\sum_{j=1}^{n} s_j Y_j\right) = \frac{\partial J}{\partial Y_i} \Delta Y_i^p \quad i \in [1, n] \tag{7.73}$$

$$\lambda_i \delta_i \mu_i \left(\sum_{j=1}^{n} v_j Y_j \right) = \frac{\partial J}{\partial Y_i} \Delta Y_i^g \quad i \in [1, n] \tag{7.74}$$

give the relevant measures of the effects of a unit of tax rate in region i on the objective function J. We may therefore interpret (7.73) as the marginal "cost" to the private sector (p) and (7.74) as the marginal "benefits" to the public sector (g). Consequently, it is clear that $\lambda(B - A)Y$ is the net overall effect of a unit change in tax rate on the objective function, or the difference between marginal benefits and marginal costs of such a change in tax rate. If $\lambda(B - A)Y > 0$, that difference is positive (marginal benefits exceed marginal costs) and, then, the tax rate is set at the maximum level, $r = 0$. If $\lambda(B - A)Y < 0$, the decrease of the objective function accompanying the losses of private investment funds produced per unit of tax rate more than offsets the increase caused by the corresponding gains due to investment in the public sector (marginal costs exceed marginal benefits) and, thus, the tax rate is set at the minimum level $r = 0$.

Finally, if $\lambda(B - A)Y = 0$, opposing private and public effects on the objective function caused by the income tax completely offset each other. The gradient of the Hamiltonian with respect th the tax rate, $H_r = \lambda(B - A)Y$, does not depend on r since H is linear in r; therefore, if $H_r \neq 0$, the tax rate corresponds to the appropriate end point of its range of feasible values, as we have seen. However, when $H_r = 0$ we can make no such simple conclusion regarding r; in addition, since H_r does not depend on r in this case, we do not have any immediate condition to find r, and we have to examine the different derivatives of H_r with respect to time. These remarks, of course, suggest interpretation of singular arcs as the case where marginal benefits equal marginal costs of the tax/investment policy.

7.2.2 Tax Rate and Public Investment as Controls

We next consider another specialization of the general model with the following characteristics: (a) transportation costs will be assumed to be zero, (b) different savings ratios will be considered for each sector in each region, (c) the objective function will be a vector function whose first component will be the sum of incomes at the terminal time t_f and whose second component will be the present value of total income, and (d) the controls considered will be the tax rate $r(t)$ and the allocations of public funds to different regions $\mu_i(t)$, where $i \in [1, n]$. The precise statement of this model, henceforth called $P2$ for short, is as follows:

$$\max \left[\sum_{i=1}^{n} Y_i(t_f); \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} Y_i(t) dt \right] \tag{7.75}$$

subject to

$$\frac{dY_i}{dt} = (1 - r(t))\psi_i \sigma_i \sum_j s_j Y_j + r(t)\mu_i(t)\delta_i \sum_j v_j Y_j, \; i \in [1, n] \tag{7.76}$$

$$Y_i(t_0) = Y_i^0 \qquad i \in [1, n] \tag{7.77}$$

$$0 \leq r(t) \leq \theta \tag{7.78}$$

$$0 \leq \mu_i(t) \qquad i \in [1, n] \tag{7.79}$$

$$\sum_{i=1}^{n} \mu_i(t) = 1 \tag{7.80}$$

In order to solve problem $P2$, we weight both components of the objective function with nonnegative weights w_1 and w_2, such that $w_1 + w_2 = 1$, and consider the following problem:

$$\max J = w_1 \sum_{i=1}^{n} Y_i(t_f) + w_2 \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} Y_i(t) dt \tag{7.81}$$

subject to the constraints (7.76), (7.77), (7.78), (7.79), and (7.80) for a range of possible weights (w_1, w_2). Solution of (7.81), subject to appropriate constraints for several different values of the weights, will of course generate an approximation to the noninferior (nondominated or Pareto optimal) set of solutions.

The application of the maximum principle is facilitated by defining the new control variables

$$q_i(t) = r(t)\mu_i(t) \qquad i \in [1, n] \tag{7.82}$$

where

$$\sum_{i=1}^{n} \mu_i(t) = 1$$

Therefore, we have

$$r(t) = \sum_i q_i(t) \quad \text{and} \quad \mu_i(t) = q_i(t)/r(t) \quad \text{for} \quad i \in [1, n]$$

Moreover, since $\sum_{i=1}^{n} \mu_i(t) = 1$ uniquely determines μ_i as a function of μ_j for $j = 1, \ldots, i-1, i+1, .., n$, the number of truly distinct controls has not changed; thus, our problem becomes

$$\max J = w_1 \sum_{i=1}^{n} Y_i(t_f) + w_2 \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} Y_i(t) dt \tag{7.83}$$

subject, for all $i \in [1, n]$, to the following constraints:

$$\frac{dY_i}{dt} = \psi_i \sigma_i \left(1 - \sum_{k=1}^{n} q_k\right)\left(\sum_{j=1}^{n} s_j Y_j\right) + q_i \delta_i \left(\sum_{j=1}^{n} v_j Y_j\right) \tag{7.84}$$

$$Y_i(t_0) = Y_i^0 \geq 0 \tag{7.85}$$

$$0 \leq q_i(t) \quad (\xi_i) \tag{7.86}$$

$$\sum_{i=1}^{n} q_i(t) \leq \theta \quad (\eta) \tag{7.87}$$

The Hamiltonian for the problem defined by (7.83) and (7.84) through (7.87) is

$$H = w_2 \exp(-\rho t) \sum_{i=1}^{n} Y_i + \left(\sum_{i=1}^{n} \lambda_i \psi_i \sigma_i\right)\left(1 - \sum_{k=1}^{n} q_k\right)\left(\sum_{j=1}^{n} s_j Y_j\right) + \left(\sum_{i=1}^{n} \lambda_i q_i \delta_i\right)\left(\sum_{j=1}^{n} v_j Y_j\right) \tag{7.88}$$

where the λ_i are adjoint variables such that the usual adjoint equations are satisfied; that is

$$-\frac{d\lambda_i}{dt} = \frac{\partial H}{\partial Y_i} = w_2 \exp(-\rho t) + s_i \left(\sum_{i=1}^{n} \lambda_i \psi_i \sigma_i\right)\left(1 - \sum_{k=1}^{n} q_k\right) + v_i \left(\sum_{i=1}^{n} \lambda_i q_i \delta_i\right) \tag{7.89}$$

where of course, since the terminal time t_f is fixed,

$$\lambda_i(t_f) = w_1 \quad \text{for} \quad i \in [1, n] \quad \text{and} \quad H(t_f) \ \text{free} \tag{7.90}$$

The maximum principle requires that we solve

$$\max_{q_1, \ldots, q_n} H \quad \text{subject to} \quad 0 \leq q_i \quad \text{and} \quad \sum_i q_i \leq \theta \qquad i \in [1, n] \tag{7.91}$$

The part of H that depends upon q_i is given by

$$-\left(\sum_{j=1}^{n} q_j\right)\left(\sum_{j=1}^{n} \lambda_j \psi_j \sigma_j\right)\left(\sum_{j=1}^{n} s_j Y_j\right) + \left(\sum_{j=1}^{n} \lambda_j q_j \delta_j\right)\left(\sum_{j=1}^{n} v_j Y_j\right) \tag{7.92}$$

We can thus construct the Lagrangian

$$L = -\left(\sum_{j=1}^{n} q_j\right)\left(\sum_{j=1}^{n} \lambda_j \psi_j \sigma_j\right)\left(\sum_{j=1}^{n} s_j Y_j\right) + \left(\sum_{j=1}^{n} \lambda_j q_j \delta_j\right)\left(\sum_{j=1}^{n} v_j Y_j\right) + \sum_{j=1}^{n} \xi_j q_j + \eta\left(\theta - \sum_{j=1}^{n} q_j\right) \quad (7.93)$$

Clearly, the first-order conditions are

$$\frac{\partial L}{\partial q_i} = -\left(\sum_{j=1}^{n} s_j Y_j\right)\left(\sum_{j=1}^{n} \lambda_j \psi_j \sigma_j\right) + \left(\sum_{j=1}^{n} v_j Y_j\right)(\lambda_i \delta_i) + \xi_i - \eta = 0 \quad i \in [1, n] \quad (7.94)$$

This expression involves terms that can be associated with the amount of savings lost in the private sector per unit of tax and the amount of savings generated in the public sector per unit of tax. In particular, these are respectively given by

$$S = \sum_{j=1}^{n} s_j Y_j \quad \text{and} \quad V = \sum_{j=1}^{n} v_j Y_j \quad (7.95)$$

Nonnegativity and complementary slackness conditions are

$$\xi_i \geq 0 \quad i \in [1, n] \quad (7.96)$$

$$q_i \geq 0 \quad i \in [1, n] \quad (7.97)$$

$$\xi_i q_i = 0 \quad i \in [1, n] \quad (7.98)$$

$$\eta \geq 0 \quad i \in [1, n] \quad (7.99)$$

$$\theta \geq \sum_{i=1}^{n} q_i \quad (7.100)$$

$$\eta\left(\theta - \sum_{i=1}^{n} q_i\right) = 0 \quad (7.101)$$

where the ξ_i and η are dual variables associated with the inequality constraints (7.86) and (7.87). Now let us define

$$\lambda_{i*}\delta_{i*} = \max\{\lambda_i \delta_i : i \in [1, n]\} \quad (7.102)$$

which may be interpreted as the greatest possible enhancement of the objective function per unit of public capital invested; hence i^* is an efficient region from the point of view of public investment. Then we have

$$\xi_k = \xi_{i^*} + V(\lambda_{i^*}\delta_{i^*} - \lambda_k\delta_k) \tag{7.103}$$

It follows that

$$\lambda_k\delta_k < \lambda_{i^*}\delta_{i^*} \Longrightarrow \xi_k > \xi_{i^*} \geq 0 \Longrightarrow \xi_k > 0 \Longrightarrow q_k = 0 \tag{7.104}$$

However, since $\lambda_i = \partial J/\partial Y_i$, this last condition is equivalent to

$$\frac{\partial J}{\partial(Y_k/\delta_k)} < \frac{\partial J}{\partial(Y_{i^*}/\delta_{i^*})} \Longrightarrow q_k = 0 \tag{7.105}$$

That is, if the change in J due to one unit of further investment in public sector k is smaller than that obtained from investing in the most efficient sector, then according to (7.105) we should not invest. Since our model derives from the relationship $Y = \sigma K$, we know that Y_i/δ_i is the equivalent capital stock that in the hands of the government would produce the current output of region i; let us call this K_i^G. Condition (7.105) may therefore be restated as

$$\frac{\partial J}{\partial K_k^G} < \frac{\partial J}{\partial K_i^G} \Longrightarrow q_k = 0 \tag{7.106}$$

To determine the extremal controls, we note from (7.94) and (7.95) that

$$V\lambda_{i^*}\delta_{i^*} - S\sum_j \lambda_j\sigma_j\psi_j = \eta - \xi_{i^*} \tag{7.107}$$

It follows that

$$V\lambda_{i^*}\delta_{i^*} - S\sum_j \lambda_j\sigma_j\psi_j \begin{cases} > 0 \Longrightarrow \ \eta - \xi_{i^*} > 0 \Longrightarrow \ \eta > \xi_j \geq 0 \Longrightarrow \eta > 0 \\ < 0 \Longrightarrow \eta - \xi_{i^*} < 0 \Longrightarrow 0 \leq \eta < \xi_{i^*} \Longrightarrow \xi_{i^*} > 0 \end{cases} \tag{7.108}$$

In the first case considered in (7.108) ($\eta > 0$), we have by complementary slackness that $\theta = \sum_{i=1}^n q_i$, and some of the q_i's will be positive; in addition, $r(t) = \sum_{i=1}^n q_i = \theta$. In the second case ($\xi_{i^*} > 0$) we have $q_{i^*} = 0$ and, thus, from (7.82) and (7.104) we know that $q_i = 0$ for $i \in [1, n]$ and $r(t) = 0$ so long as $\mu_i > 0$. The interpretation of this policy is straightforward. Considered per unit of tax rate, V and S defined by (7.95) are, respectively, total savings of the public sector and the total loss of savings of the private sector. When the funds V are invested in region i, they produce an output of $\delta_i V$. Since $\lambda_i = \partial J/\partial Y_i$, it is clear that $Vp_i\delta_i$ is the slope of J with respect to r when all public funds go to region i. Since $\lambda_{i^*}\delta_{i^*} = \max_i\{\lambda_i\delta_i\}$, it is also clear that $V\lambda_{i^*}\delta_{i^*}$ is the increment in J per unit of tax rate when the public

funds are invested in the best use, that is, in the region that will produce the highest increase in J. Analogously, S, the loss of savings of the private sector, may be translated into the loss of savings for the jth region, $\psi_j S$, and an income loss for the jth region, $\sigma_j \psi_j S$.This income loss has an effect $\lambda_j \sigma_j \psi_j S$ on J; the sum of such effects for all region is $S \sum \lambda_j \sigma_j \psi_j$.

We are now in a position to articulate the following rule: if the increase of J produced by the investment of additional public funds per unit of tax rate in the best available opportunity ($V\ \lambda_{i*} \delta_{i*}$) is greater than the decrease in J produced by the loss of savings in the private sector per unit of tax rate ($S \sum_j \lambda_j \sigma_j \psi_j$), taking into account the prevailing private allocation pattern ψ_i^j for $i \in [1, n]$, then tax at the maximum rate and invest in the best regions. Those best regions must belong to the set I^* where

$$I^* = \{i \in [1, n] : \lambda_i \delta_i \geq \lambda_j \delta_j \quad j \in [1, n]\} \tag{7.109}$$

Note that the following decision rules result:

1. If I^* has more than one element, then the marginal benefits of further allocation of public investment to regions in I^* are the same, and we have singularity in the optimal allocations. In this case the optimal allocations would be found using the condition
$$\theta = \sum_{i \in I^*} q_i^*$$
and vanishing of the successive derivatives with respect to time of the gradient of the Hamiltonian with respect of q_i where $i \in I^*$.
2. If the marginal benefits in the best region, $V\ \lambda_{i*} \delta_{i*}$, equal the marginal cost $S \sum_{j=1}^n \lambda_j \sigma_j \psi_j$, we may have two levels of singularity: one in the determination of the tax rate $r(t) = \sum_{i=1}^n q_i(t)$, and if I^* has more than one element, another in the determination, as above, of the $q_i, i \in I^*$, themselves.
3. Finally, if $V\ \lambda_{i*} \delta_{i*} < S \sum_{j=1}^n \lambda_j \sigma_j \psi_j$, then even the best public investment opportunity is not profitable in comparison with the given pattern of investments of the private sector; in such a case $r = 0$ and $\mu_i = 0$ for $i \in [1, n]$.

The existence of and properties of singular solutions, as has been demonstrated above, are necessary to fully specify an optimal policy. The interpretation of the singular policies involves a comparison of appropriate marginal benefits for both the public and private sectors. In that singular solutions may exist over finite portions of the planning horizon, failure to include singular controls in the specification of the optimal policy will correspond to an incomplete analysis. This has not been emphasized in most of the published regional economic growth theory literature.

7.2.3 Equal Public and Private Savings Ratios

As we indicated in the presentation of the general model and its relationship to other models, the special case of identical public and private savings ratios has played a

central role in the models reported in the published literature. For this reason we devote our attention in this section to the analysis of that special case. To make the analysis as simple as possible we consider the case of only the tax rate as a control variable; the public investment shares will be assumed to be exogenously determined. Following this we will consider the more general case of both tax rate and public investment shares as controls.

When we take $s_j = v_j$, the general model [namely (7.19), (7.20), (7.21), and (7.23)] is substantially simplified. In fact, for the case of two regions, objective function weights w_1 and w_2 and considering only the tax rate as a control, we obtain the following formulation, called $P3$:

$$\max J = w_1 cx(t_f) + w_2 \int_{t_0}^{t_f} \exp(-\rho t) cx(t) \tag{7.110}$$

subject to

$$\begin{aligned} \dot{x} &= Ax(t) + r(t)(B - A)x(t) \qquad (7.111)\\ x(t_0) &= x^0 \Longrightarrow x_i(t_0) = x_i^0 \\ 0 &\le r(t) \le \theta \quad t \in [t_0, t_f] \end{aligned}$$

The transformation used to specify (7.110) and (7.111) is as follows:

$$x = (x_1, \ldots, x_n)^T \tag{7.112}$$

where

$$x_i = s_i Y_i \quad i \in [1, n] \tag{7.113}$$

and

$$\begin{aligned} A &= (a, a, \ldots, a) \qquad (7.114)\\ a &= (s_1\psi_1\sigma_1, \ldots, s_n\psi_n\sigma_n)^T \\ B &= (b, b, \ldots, b) \\ b &= (s_1\mu_1\delta_1, \ldots, s_n\mu_n\delta_n)^T \\ c &= \left(\frac{1}{s_1}, \frac{1}{s_2}, \ldots, \frac{1}{s_n}\right) \end{aligned}$$

All other notation is as used previously. In order to apply Pontryagin's maximum principle we need to construct the Hamiltonian which for the present case takes on the form

$$H = w_2 \exp(-\rho t) cx + \lambda[Ax + r(B - A)x] \tag{7.115}$$

where λ is the vector of adjoint variables. To determine the maximum of the Hamiltonian with respect to the control $r(t)$ we must specify the so-called switching function which is the gradient of H with respect to r, that is $H_r = \lambda(B - A)x$.

Thus, the extremal tax rate is

$$r^* = \begin{cases} \theta & \text{if } \lambda(B-A)x > 0 \\ 0 & \text{if } \lambda(B-A)x < 0 \\ \text{undetermined} & \text{if } \lambda(B-A) = 0 \end{cases} \tag{7.116}$$

Clearly, we must specify the adjoint equation in order to analyze further the extremal policy (7.116). The adjoint equation is

$$-\frac{d\lambda}{dt} = H_x = \lambda[A + r(B-A)] + w_2 \exp(-\rho t)c \tag{7.117}$$

with the terminal conditions

$$\lambda(t_f) = w_1 c \tag{7.118}$$

that are recognized as the transversality conditions.

The standard procedure for determining singular controls is to take successive time derivatives of the switching function or gradient of the Hamiltonian, H_r in this case. Since H is linear in the control $r(t)$, H_r is independent of $r(t)$. One can proceed by taking successive time derivatives of H_r and making appropriate substitutions using the adjoint equations and the dynamics of the problem in an attempt to find an explicit expression for $r(t)$ on a singular arc. In the present case we may show that this procedure fails to yield an explicit expression for $r(t)$ along a singular arc of time (ultimately reducing to a trivial identity) and, so, we conclude that a singular control does not arise. To follow this paradigm, we first note that the switching function can be written as

$$\begin{aligned} \lambda(B-A)x &= \lambda(b-a, \ldots, b-a)x \\ &= \lambda(b-a)\left(1^T x\right) \end{aligned}$$

However, since

$$1^T \cdot x = \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} s_i Y_i > 0$$

in all cases, we must have $\lambda(b-a) = 0$, which implies that

$$\lambda(B-A) = 0 \tag{7.119}$$

Expression (7.119) therefore becomes a necessary condition for a singular arc. If we now recall expression (7.117), it is clear that on a singular arc the adjoint equation is

$$-\frac{d\lambda}{dt} = \lambda A + w_2 \exp(-\rho t)c \tag{7.120}$$

By inspection we have that the adjoint variables do not depend on r, our control variable. Thus, successive time derivatives of (7.119) combined with substitutions based on (7.120) will never yield an expression for r, which indicates that a singular arc for the problem as presently considered does not exist.

Let us now consider the case of controls involving not only $r(t)$, the income tax rate, but also the $\mu_i(t)$, the public investment shares. Moreover, we assume that the public and private savings ratios are unequal. In that case it may be shown that the condition for a singular arc is

$$V\lambda_{i*}\delta_{i*} = S\sum_{j=1}^{n}\lambda_j\sigma_j\psi_j \tag{7.121}$$

where, for convenience, we recall that

$$\begin{aligned} V &= \sum_{i=1}^{n} v_i Y_i \\ \lambda_{i*}\delta_{i*} &= \max\{\lambda_i\delta_i : i \in [1,n]\} \\ S &= \sum_i s_i Y_i \end{aligned}$$

When $s_i = v_i$ expression (7.121) may be simplified to

$$\lambda_{i*}\delta_{i*} = \sum_{j=1}^{n}\lambda_j\sigma_j\psi_j \tag{7.122}$$

We may assume that when (7.122) obtains for a finite period of time it does so in such a way that i^* does not change over that period. In that case

$$\frac{d\lambda_{i*}}{dt}\delta_{i*} = \sum_{j=1}^{n}\frac{d\lambda_j}{dt}\sigma_j\psi_j \tag{7.123}$$

The adjoint equations for the problem with both types of control variables (tax rate and public investment shares) are

$$-\frac{d\lambda}{dt} = w_2\exp(-\rho t) + s_i\left(\sum_{j=1}^{n}\lambda_j\sigma_j\psi_j\right)\left(1-\sum_{j=1}^{n}q_j\right) + v_i\left(\sum_{i=1}^{n}\lambda_i\delta_i q_i\right) \tag{7.124}$$

In addition, we know that

$$\lambda_i\delta_i < \lambda_{i*}\delta_{i*} \Longrightarrow q_i = 0 \tag{7.125}$$

and consequently

$$\sum_{i=1}^{n}\lambda_i\delta_i q_i = \lambda_{i*}\delta_{i*}\sum_{i=1}^{n}q_i \tag{7.126}$$

Thus, the adjoint equations on a singular path become

$$-\frac{d\lambda_j}{dt} = w_2 \exp(-\rho t) + s_j \lambda_{i*} \delta_{i*} \tag{7.127}$$

By making appropriate substitutions into (7.122) and (7.123), we obtain the following expressions:

$$\lambda_{i*} = \frac{w_2 \exp(-\rho t) \left(\sum_{j=1}^{n} \sigma_j \psi_j - \delta_{i*} \right)}{\delta_{i*} \left(\delta_{i*} s_{i*} - \sum_{j=1}^{n} \sigma_j \psi_j s_j \right)} \tag{7.128}$$

$$\frac{d\lambda_{i*}}{dt} = \frac{\rho w_2 \exp(-\rho t) \left(\sum_j \sigma_j \psi_j - \delta_{i*} \right)}{\delta_{i*} \left(\delta_{i*} s_{i*} - \sum_{j=1}^{n} \sigma_j \psi_j s_j \right)} \tag{7.129}$$

By substituting both (7.128) and (7.129) into (7.127), the adjoint equations become

$$\rho \left(\sum_{j=1}^{n} \sigma_j \psi_j - \delta_{i*} \right) = \delta_{i*} \left(s_{i*} \sum_{j=1}^{n} \sigma_j \psi_j - \sum_{j=1}^{n} \sigma_j \psi_j s_j \right) \tag{7.130}$$

which is a necessary condition for the existence of singular arcs. This necessary condition only involves structural parameters of the problem and will not, in general, be satisfied for any i^*. Expression (7.122), (7.123), and (7.127) do not depend on q_i or Y_i and, consequently, imply that further differentiations with respect to time will not lead to explicit expressions for the q_i's; hence, singular arcs do not typically exist for this more general problem. Notably, the results of Sakashita (1967a) and Ohtsuki (1971) are seen *a fortiori* to be correct, although these authors did not consider singular controls.

We now will consider a special case of this same problem in which the tax rate and the investment shares are the controls and the public and private savings ratios are equal, and an additional assumption regarding the relationship of certain structural parameters is invoked. This version is of interest because it leads to an exact analytical representation of the noninferior set (set of Pareto optimal solutions). The switching function which determines whether $r(t) = 0$ or $r(t) > 0$ is the same as for the previous case, namely, expression (7.121):

$$V \lambda_{i*} \delta_{i*} - S \sum_{j=1}^{n} \lambda_j \sigma_j \psi_j \tag{7.131}$$

Under the assumption of an equal, normalized savings ratio for both the public and private sectors of each region

$$V = S = 1$$

and expression (7.121) may be simplified to

$$\lambda_{i*}\delta_{i*} = \sum_{j=1}^{n} \lambda_j \sigma_j \psi_j \tag{7.132}$$

From the transversality conditions $\lambda_i(t_f) = w_1$ for $i \in [1, n]$; hence, the value of the switching function at $t = t_f$ is

$$\delta_{i*} - \sum_{j=1}^{n} \sigma_j \psi_j \tag{7.133}$$

where $w_1 = 1$ has been normalized as well. Let us assume that $\delta_i - \sum_j \sigma_j \psi_j < 0$ for all $i \in [1, n]$; then $q_i = 0$ for all $i \in [1, n]$ and $r = 0$. In the neighborhood of t_f, the adjoint equations become

$$-\frac{d\lambda_i}{dt} = w_2 \exp(-\rho t) + s_i \lambda_i \sum_{j=1}^{n} \lambda_j \sigma_j \psi_j \tag{7.134}$$

Expression (7.134) will be used to analyze the behavior of the switching function to the left of $t = t_f$.

The derivative of the switching function with respect to time is

$$\frac{d}{dt}\left(\lambda_i \delta_i - \sum_{j=1}^{n} \lambda_j \sigma_j \psi_j\right) = \frac{d\lambda_i}{dt}\delta_i - \sum_{j=1}^{n} \dot{\lambda}_i \sigma_j \psi_j \tag{7.135}$$

Using the adjoint equations (7.134), we obtain

$$\begin{aligned} &\frac{d}{dt}\left(\lambda_i \delta_i - \sum_{j=1}^{n} \lambda_j \sigma_j \psi_j\right) \\ &= w_2 \exp(-\rho t)\left(\sum_{j=1}^{n} \sigma_j \psi_j - \delta_i\right) \sum_k \lambda_k \sigma_k \psi_k \left(\sum_{j=1}^{n} s_j \sigma_j \psi_j - s_i \delta_i\right) \end{aligned} \tag{7.136}$$

By assumption $\sum_{j=1}^{n} \sigma_j \psi_j - \delta_i > 0$ for all i, and it is also clear that $w_2 \exp(-\rho t) > 0$ for all t. Thus, the first term of the right-hand side of (7.136) is clearly nonnegative for all $t \leq t_f$. From the transversality conditions, we have

$$\sum_{j=1}^{n} \lambda_j \sigma_j \psi_j|_{t=t_f} = w_1 \sum_{j=1}^{n} \sigma_j \psi_j > 0 \tag{7.137}$$

In addition, by using the adjoint equation we obtain

$$-\sum_{j=1}^{n}\frac{d\lambda_i}{dt}\sigma_j\psi_j = w_2\exp(-\rho t)\sum_{j=1}^{n}\sigma_j\psi_j + \sum_{i=1}^{n}\lambda_i\sigma_i\psi_i\sum_{j=1}^{n}s_j\sigma_j\psi_j \quad (7.138)$$

It is clear then that $\sum_{i=1}^{n}\lambda_i\sigma_i\psi_i > 0$ for all $t \le t_f$ because of the exponential nature of solution to (7.138). Thus, if we assume that

$$\sum_{j=1}^{n}s_j\sigma_j\psi_j > s_i\delta_i \quad (i = 1,\ldots,n) \quad (7.139)$$

It follows immediately that

$$\frac{d}{dt}\left(\lambda_i\delta_i - \sum_{j=1}^{n}\lambda_j\sigma_j\psi_j\right) > 0 \quad i \in [1,n] \quad t \le t_f \quad (7.140)$$

Then by assumption

$$\left(\lambda_i\delta_i - \sum_{j=1}^{n}\lambda_j\sigma_j\psi_j\right)\Bigg|_{t=t_f} = w_1\left(\delta_i - \sum_{j=1}^{n}\sigma_j\psi_j\right) < 0 \quad i \in [1,n]$$

It is now evident that the switching function will be negative over the whole planning period and thus $r^*(t) = 0$ for all $t \in [t_0, t_f]$. In this particular case the optimal policy $r^*(t) = 0$ for all $t \in [t_0, t_f]$ does not depend on w_1 or w_2, and, therefore, the noninferior set will consist of the points on the straight line segment joining $(J_1^*, 0)$ and $(0, J_2^*)$, where

$$J_1^* = \sum_i Y_i^*(t_f) \quad J_2^* = \int_{t_0}^{t_f}\exp(-\rho t)\sum_i Y_i^*(t)dt \quad (7.141)$$

when the Y_i^* denote state functions correspond to the optimal policy. The computation of J_1^* and J_2^* is straightforward and is not reported here.

7.2.4 *Sufficiency*

The conditions established by Pontryagin's maximum principle are only necessary conditions, and therefore any solution derived from them is only a candidate for optimality. In this section, for the Harrod-Domar type regional growth models presented above, we will analyze the circumstances which cause their necessary conditions to be sufficient. We will anlyze sufficiency using the Arrow-Kurz sufficiency theorem

presented in Chapter 3. In particular, the Arrow-Kurz sufficiency theorem, when applied to the class of optimal regional growth models we have considered, tells us that a policy $[Y^*(t), Z^*(t)]$ obtained from the maximum principle, where $Y^*(t)$ is the vector of regional incomes and $Z^*(t)$ is the vector of controls [which in some cases may have only one component, the tax rate $r(t)$], will produce a global maximum of the objective function $J(Z(t))$ if $H^*(Y, \lambda, t)$ is concave in the state variables Y for all $t \in [t_0, t_f]$, where

$$H^*(Y, \lambda, t) = \max_{Z \in \Omega} H(Y, \lambda, Z, t)$$

and $H(Y, \lambda, Z, t)$ is the Hamiltonian. The specific problem we wish to analyze is the general regional investment allocation model introduced in Section 7.2.2, namely

$$\max J(r, \mu_1, \ldots, \mu_n) = w_1 \sum_{i=1}^{n} Y_i(t_f) + w_2 \int_{t_0}^{t_f} \exp(-\rho t) \sum_{i=1}^{n} Y_i(t) dt \quad (7.142)$$

subject to

$$\frac{dY_i}{dt} = (1 - r(t))\psi_i \sigma_i \sum_j s_j Y_j + r(t)\mu_i(t)\delta_i \sum_j v_j Y_j \quad i \in [1, n] \quad (7.143)$$

$$Y_i(t_0) = Y_i^0 \qquad i \in [1, n] \quad (7.144)$$

$$0 \leq r(t) \leq \theta \quad (7.145)$$

$$0 \leq \mu_i(t) \qquad i \in [1, n] \quad (7.146)$$

$$\sum_{i=1}^{n} \mu_i(t) = 1 \quad (7.147)$$

This is a Bolza-type problem, since the criterion values of the state variables at one of the end points ($t = t_f$) of the interval $[t_0, t_f]$.

In order to apply the Arrow-Kurz theorem, it is necessary to have a problem whose objective function depends only on the path followed. Using the transformation suggested by Hadley and Kemp (1971), it is possible to change the objective function so as to obtain a problem of Lagrange type, i.e., one whose objective function does not depend on the values of the state variables and/or the control variables at the end points of $[t_0, t_f]$. To this end let us introduce a new state variable $Y_{n+1}(t)$ and consider the problem

$$\max J_0(r(t), \mu_1(t), \ldots, \mu_n(t)) = \int_{t_0}^{t_f} \left(Y_{n+1}(t) + w_2 \exp(-\rho t) \sum_i Y_i(t) \right) dt \quad (7.148)$$

subject, for all $i \in [1, n]$, to the following constraints:

$$\frac{dY_i}{dt} = (1 - r(t))\psi_i \sigma_i \sum_j s_j Y_j + r(t)\mu_i(t)\delta_i \sum_j v_j Y_j \tag{7.149}$$

$$\frac{dY_{n+1}}{dt} = 0 \tag{7.150}$$

$$Y_i(t_0) = Y_i^0 \tag{7.151}$$

$$0 = w_1 \sum_i Y_i(t_f) - t_f Y_{n+1}(t_f) \tag{7.152}$$

$$0 \leq r(t) \leq \theta \tag{7.153}$$

$$0 \leq \mu_i(t) \tag{7.154}$$

$$\sum_{i=1}^{n} \mu_i(t) = 1 \tag{7.155}$$

That this second formulation of the problem is equivalent to the first is easy to show, and so we do not further elaborate on it. Note that the Hamiltonian for the transformed problem is

$$\begin{aligned} H(Y, \lambda, r, \mu, t) &= Y_{n+1} + \exp(-\rho t) \sum_{i=1}^{n} Y_i \\ &\quad + \sum_{i=1}^{n} \lambda_i \left[(1 - r)\psi_i \sigma_i \sum_{j=1}^{n} s_j Y_j + r\mu_i \delta_i \sum_{j=1}^{n} v_j Y_j \right] + \lambda_{n+1} \cdot 0 \\ &= Y_{n+1} + \exp(-\rho t) \sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} \lambda_i \psi_i \sigma_i \sum_{j=1}^{n} s_j Y_j \\ &\quad - r \sum_{i=1}^{n} \lambda_i \psi_i \sigma_i \sum_{j=1}^{n} s_j Y_j + r \sum_{i=1}^{n} \lambda_i \mu_i \delta_i \sum_{j=1}^{n} v_j Y_j \end{aligned} \tag{7.156}$$

We first analyze the case in which $r(t)$ is the only control variable; in that case the μ_i's are considered to be constants. The switching function is the same as that found previously, namely

$$F_s = \sum_{i=1}^{n} \lambda_i \mu_i \delta_i \sum_{j=1}^{n} v_j Y_j - \sum_{i=1}^{n} \lambda_i \psi_i \sigma_i \sum_{j=1}^{n} s_j Y_j$$

and gives rise to three cases

$$F_s \begin{cases} > 0 \Longrightarrow r = \theta \\ < 0 \Longrightarrow r = 0 \\ = 0 \Longrightarrow \text{singular control } r \in [0, \theta] \end{cases}$$

The forms of the Hamiltonian corresponding to these cases are

$$F_s > 0 : H^*(Y,\lambda,t) = Y_{n+1} + \exp(-\rho t)\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n}\lambda_i\psi_i\sigma_i\sum_{j=1}^{n} s_j Y_j \quad (7.157)$$
$$- \theta\left[\sum_{i=1}^{n}\lambda_i\psi_i\sigma_i\sum_{j=1}^{n} s_j Y_j - \sum_{i=1}^{n}\lambda_i\mu_i\delta_i\sum_{j=1}^{n} v_j Y_j\right]$$
$$F_s \leq 0 : H^*(Y,\lambda,t) = Y_{n+1} + \exp(-\rho t)\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n}\lambda_i\psi_i\sigma_i\sum_{j=1}^{n} s_j Y_j \quad (7.158)$$

Both expressions are linear and, thus, concave in the state variables and, therefore, the sufficiency theorem holds.

We next consider as control variables $r(t)$ and $\mu_1(t),\ldots,\mu_n(t)$. As before, we introduce the control vector

$$q = [q_1(t),\ldots,q_n(t)]^T$$

defined by

$$q_i(t) = r(t)\mu_i(t) \quad i \in [1,n] \quad t \in \left[t_0, t_f\right])$$

The Hamiltonian then becomes

$$\begin{aligned} H(Y,\lambda,r,\mu,t) &= Y_{n+1} + \exp(-\rho t)\sum_{i=1}^{n} Y_i \\ &\quad + \sum_{i=1}^{n}\lambda_i\left[(1-r)\psi_i\sigma_i\sum_{j=1}^{n} s_j Y_j + r\mu_i\delta_i\sum_{j=1}^{n} v_j Y_j\right] \\ &= Y_{n+1} + \exp(-\rho t)\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n}\lambda_i\psi_i\sigma_i\sum_{j=1}^{n} s_j Y_j \\ &\quad - \sum_{k=1}^{n} q_k\sum_{i=1}^{n}\lambda_i\psi_i\sigma_i\sum_{j=1}^{n} s_j Y_j \\ &\quad + \sum_{i=1}^{n}\lambda_i q_i\delta_i\sum_{j=1}^{n} v_j Y_j \end{aligned} \quad (7.159)$$

and the set of admissible values for the optimal controls is

$$\Omega = \left\{q : \sum_{i=1}^{n} q_i \leq \theta \quad q_i \geq 0\right\}$$

The switching function is

$$\lambda_k \delta_k \sum_{j=1}^{n} v_j Y_j - \sum_{i=1}^{n} \lambda_i \psi_i \sigma_i \sum_{j=1}^{n} s_j Y_j$$

It is left as an exercise for the reader to determine that the Hamiltonian is linear and thus concave when evaluated for the optimal control policy, making the necessary conditions also sufficient.

7.3 A Computable Theory of Regional Public Investment Allocation

In this section, we want to consider a model for optimal regional allocation of public investment that corrects several of the shortcomings of the simple models considered in Section 7.2. In particular, we will introduce a model that possesses the following characteristics:

1. growth dynamics are not based on a constant returns assumption and allow increasing returns to scale;
2. the production technology employs both capital and labor inputs;
3. capital markets are in equilibrium;
4. private and public capital, the latter allowing infrastructure investment decisions to be modeled, are distinguished from one another;
5. population evolves over time in accordance with a Hotelling-type of diffusion model that includes births, deaths, and location-specific ecological carrying capacities;
6. capital augmenting technological change is allowed and is endogenous in nature;
7. regulatory and fiscal policy constraints may be imposed; and
8. the optimization criterion is the present value of the national income time stream.

The model relies on the same spatial disaggregation of macroeconomic identities relating the rate of change of capital stocks to investments and depreciation that is characteristic of Datta-Chaudhuri (1967), Sakashita (1967b), Ohtsuki (1971), Domazlicky (1977), Bagchi (1984) and Friesz and Luque (1987) and was employed in Section 7.2 of this chapter. The model introduced below, however, differs from historical regional growth models in that it does not rely on the assumption of a constant proportionate rate of labor force growth for each region. The constant proportionate growth (CPG) model of labor and population allows, as was seen in Section 7.1.3, the dynamics for population growth to be uncoupled from those of capital formation. Thus, in CPG models, population always grows exponentially with respect to time and shows no response to changes in population density or regional income. In the presentation below, we replace the unrealistic CPG model of

population growth with a Hotelling-type model that includes the effects of spatial diffusion and of ecological carrying capacities of individual regions and is intrinsically coupled to the dynamics of capital formation.

7.3.1 The Dynamics of Capital Formation

Once again we employ basic macroeconomic identities to describe the relationship of the rate of change of capital to investment, output and savings. In particular we take output to be a function of capital and labor, and the rate of change of capital is equated to investment less any depreciation of capital that may occur. That is:

$$\frac{dK}{dt} = I - \delta K \tag{7.160}$$

where K is capital, dK/dt is the time rate of change of capital, I is investment, and δ is now an abstract depreciation rate (rather than an output-capital ratio as in Section 7.2). Subsequently δ_p will be the depreciation rate of private capital and δ_g the depreciation rate of public capital. Of course, (7.160) is an aspatial model. It is also important to recognize that (7.160) is an equilibrium model for which the supply of capital is exactly balanced against the demand for capital. We shall maintain this assumption of capital market equilibrium throughout the development that follows. It is also worth noting that (7.160) is the foundation of the much respected work by Arrow and Kurz (1970) exploring the interdependence of aspatial private and public sector growth dynamics. To spatialize (7.160) as well as to introduce a distinction between the public and private sectors we write:

$$\frac{dK_i^p}{dt} = I_i^p - \delta_p K_i^p \quad \text{and} \quad \frac{dK_i^g}{dt} = I_i^g - \delta_g K_i^g \tag{7.161}$$

where the subscript $i \in [1, n]$ refers to the ith of n regions and the superscripts p and g refer to the private and public (governmental) sectors respectively.

Further detail can be introduced into the above dynamics by defining c_i to be the consumption rate of region i and r to be a tax rate imposed by the central government on each region's output. We also now define ω_i to be the share of tax revenues allocated to subsidize private investments in region i, and ν_i to be the share of tax revenues allocated to public (infrastructure) investments in region i. Also Y_i will be the output of region i. To keep the presentation simple, we assume that all capital (private as well as public) is immobile, although this assumption can be relaxed at the expense of more complicated notation. Consequently, the following two identities hold for all $i \in [1, n]$:

$$I_i^p = (1 - c_i - r)\, Y_i + \omega_i r \sum_{j=1}^{n} Y_j \quad \text{and} \quad I_i^g = \nu_i r \sum_{j=1}^{n} Y_j \tag{7.162}$$

By virtue of the definitions of ω_i and ν_i the following constraints obtain:

$$\sum_{i=1}^{n} (\omega_i + \nu_i) = 1 \tag{7.163}$$

$$0 \leq \omega_i \leq 1 \quad \text{and} \quad 0 \leq \nu_i \leq 1 \tag{7.164}$$

which require that allocations cannot exceed the tax revenues collected and must be nonnegative[1]. We further assume that the ith region's intrinsic technology, ignoring for the moment technological innovation, is described by a production function of the form

$$Y_i = F_i \left(K_i^p, K_i^g, L_i\right) \tag{7.165}$$

where L_i is the labor force (population) of the i^{th} region.

It follows at once from (7.161) and (7.165) that the dynamics for the evolution of private and public sector capital are

$$\frac{dK_i^p}{dt} = (1 - c_i - r)\, F_i \left(K_i^p, K_i^g, L_i\right) + \omega_i r \sum_{j=1}^{n} F_j \left(K_j^p, K_j^g, L_j\right) - \delta_p K_i^p$$

$$\forall i \in [1, n] \tag{7.166}$$

$$\frac{dK_i^g}{dt} = \nu_i r \sum_{j=1}^{n} F_j \left(K_j^p, K_j^g, L_j\right) - \delta_g K_i^g \qquad \forall i \in [1, n] \tag{7.167}$$

It is important to note that we have made no assumption regarding constant returns to scale in articulating the above dynamics.

7.3.2 Population Dynamics

Traditionally, the literature on neoclassical economic growth, as we noted above, has assumed a constant proportionate rate of labor force growth. As population and labor force are typically treated as synonymous, this means that models from this literature employ quite simple population growth models of the form

$$\frac{dL_i}{dt} = \pi_i \qquad L_i\,(0) = L_i^0 \tag{7.168}$$

for every region $i \in n$ where π_i is a constant. This means that population (labor force) always grows according to the exponential law $L_i\,(t) = L_i^0 e^{\pi_i t}$ regardless of any other assumptions employed. The CPG assumption is decidedly unrealistic,

[1] Other tax schemes, such as own-region taxes, can easily be described. The one chosen here is meant to be illustrative.

limits the policy usefulness of economic growth models based on it and calls out to be replaced with a richer model of population and labor force change over time and space.

We replace (7.168) with a spatial diffusion model of the Hotelling-type.[2] In Hotelling-type population models, migration is based on the noneconomic notion of diffusion wherein populations seek spatial niches that have been previously unoccupied. This means that unlike (7.168) population will not become inexorably denser at a given point in space, but rather that population density may rise and fall over time. Yet because we will link this diffusion process to the capital formation process there will be a potential for population to concentrate where infrastructure agglomeration economies occur. Furthermore, we will employ a version of the spatial diffusion process that includes a logistic model of birth/death processes and specifically incorporates the ecological carrying capacity of each location alternative. These features will inform and be informed by the capital dynamics (7.44) and (7.45), resulting in an economic growth model that is intrinsically more realistic than would result from rote adherence to the neoclassical paradigm.

Hotelling's original model is in the form of a partial differential equation which is very difficult to solve for realistic spatial boundary conditions and is not readily coupled with ordinary differential equations such as (7.44) and (7.45). Puu (1989) and Puu (1997), however, have suggested a multiregion alternative to Hotelling's model which captures key features of the diffusion process and the birth/death process in a more tractable mathematical framework. Specifically, Puu (1989) proposes, if the population of region i is denoted as P_i, the following dynamics:

$$\frac{dP_i}{dt} = \gamma_i P_i \left(\zeta_i - P_i\right) + \sum_{j \in [1,n] \setminus i} \kappa_j \left(P_j - P_i\right) \qquad \forall i \in [1,n] \tag{7.169}$$

where γ_i, ζ_i and κ_i are positive exogenous parameters. The idea here is that the term $\sum_{j \in [1,n] \setminus i} \kappa_j \left(P_j - P_i\right)$ is roughly analogous to diffusion in that it draws population from regions with higher population density toward regions with lower population density. Typically κ_j is referred to as the *coefficient of diffusion* for region $j \in n$. The entity ζ_i is sometimes called the *fitness measure* and describes the *ecological carrying capacity* of region $i \in n$; its units are population. The parameter γ_i ensures dimensional consistency and has the units of $(time)^{-1}$. Clearly this model is not equivalent to Hotelling's, but it does capture the essential ideas behind diffusion-based population growth and migration and is substantially more tractable from a computational point of view since (7.169) is a system of ordinary (as opposed to partial) differential equations.

Moreover, the population dynamics (7.169) can be considerably enriched by allowing the fitness measure to be locationally and infrastructurally specific as we now show. Specifically, we postulate that

$$\zeta_i = V_i \left(K_i^g, t\right) + \Psi_i \left(t\right) \tag{7.170}$$

[2] See Hotelling (1978).

where $V_i\left(K_i^g, t\right)$ describes the effect of infrastructure on carrying capacity and $\Psi_i(t)$ is the natural or ambient carrying capacity that exists in the absence of infrastructure investment. It is important to understand that by "carrying capacity" we mean the population that a region can sustain. As such, (7.170) expresses the often made observation that each individual region is naturally prepared to support a specific population level, and that level may vary with time and be conditioned by manmade infrastructure. Puu (1989) observes that population models like that presented above have one notable shortcoming: it is possible, for certain initial conditions, that population trajectories will include periods of negative population. Negative population is of course meaningless and population trajectories with this property cannot be accepted as realistic. Consequently, we must include in the final optimal control formulation a state space constraint that forces population to remain nonnegative.

7.3.3 Technological Change

None of the above presentation depends on the idea of balanced growth or the assumption that a long-run equilibrium exists in the usual sense. Neither do we assume that technological progress must obey some type of neutrality. So when introducing technological progress we have a free hand to explore any type of technological progress. In particular, we are not restricted to labor augmenting progress or progress that enhances overall output, and we can explore capital augmenting progress, as is natural since our interest is in the role of public capital (infrastructure). That is, we shall concentrate on technological progress which augments K_i^g. Consequently for each region i, we shall update the associated production function by making the substitution

$$K_i^g \Longrightarrow \Phi(t) K_i^g \tag{7.171}$$

where Φ is a scalar function of time describing the extent of infrastructure augmenting technological progress. Technological progress for our model is endogenously generated through separate dynamics for Φ. We postulate that public capital augmenting technological progress occurs when the national output/public-capital ratio falls below some threshold and is zero when the ratio exceeds that threshold. That is, the rate of technological progress $d\Phi/dt$ obeys

$$\frac{d\Phi}{dt} > 0 \quad \text{if} \quad \sigma \leq \Theta \tag{7.172}$$

$$\frac{d\Phi}{dt} = 0 \quad \text{if} \quad \sigma > \Theta \tag{7.173}$$

where σ is the output/public-capital ratio and $\Theta \in \Re^1_+$ is a known reference threshold that determines the need for and the fact of technological progress. Note that the output/public-capital ratio is dependent on multiple state variables:

$$\sigma\left(K^p, K^g, L, \Phi\right) = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n K_i^g} = \frac{\sum_{i=1}^n F_i\left(K_i^p, \Phi K_i^g, L_i\right)}{\sum_{i=1}^n K_i^g} \tag{7.174}$$

where K^p, K^g and L are vectors of private capital, public capital, and labor respectively. Moreover, σ is implicitly time dependent since it is constructed from time-varying entities. It follows that the rate of technological progress is

$$\frac{d\Phi}{dt} = \tau\left[\Theta - \sigma\left(K^p, K^g, L, \Phi\right)\right]_+ \tag{7.175}$$

where $\tau \in \Re^1_{++}$ is an exogenous positive constant of proportionality and $[.]_+$ is the nonnegative orthant projection operator with the property that $[Q]_+ = \max(0, Q)$ for an arbitrary argument Q.

7.3.4 Criterion Functional and Final Form of the Model

In what follows, we have assumed for simplicity that there is full labor force participation, so that $L_i = P_i$.We wish to maximize the present value of the national income time steam for the time interval $\left[t_0, t_f\right]$ expressed as

$$\max J(K^p, K^g, P) = \sum_{i=1}^n \int_{t_0}^{t_f} \exp(-\rho t)\, F_i\left(K_i^p, K_i^g, P_i\right) dt \tag{7.176}$$

where $\rho > 0$ is the constant nominal interest rate and P is presently a vector of regional specific populations:

$$P = \left(P_i : i \in n\right)$$

This maximization is to be carried out relative to the dynamics and constraints developed above. Hence, the final form of the model is

$$\max J\left(K^p, K^g, P\right)$$

subject to

$$\frac{dK_i^p}{dt} = (1 - c_i - r)\, F_i\left(K_i^p, \Phi K_i^g, P_i\right) + \omega_i r \sum_{j=1}^n F_j\left(K_j^p, \Phi K_j^g, P_j\right) - \delta_p K_i^p$$
$$\forall i \in [1, n]$$

and

$$\frac{dK_i^g}{dt} = v_i r \sum_{j=1}^{n} F_j\left(K_j^p, \Phi K_j^g, P_j\right) - \delta_g K_i^g \qquad \forall i \in [1,n]$$

$$\frac{dP_i}{dt} = \gamma_i P_i\left(\zeta_i - P_i\right) + \sum_{j \neq i} \kappa_j\left(P_j - P_i\right) \qquad \forall i \in [1,n]$$

$$\frac{d\Phi}{dt} = \tau\left[\Theta - \sigma\left(K^p, K^g, P, \Phi\right)\right]_+$$

$$\sum_{i=1}^{n}(\omega_i + v_i) = 1 \qquad \forall i \in [1,n]$$

$$0 \leq \omega_i \leq 1 \qquad \forall i \in [1,n]$$

$$0 \leq v_i \leq 1 \qquad \forall i \in [1,n]$$

$$P_i \geq 0 \qquad \forall i \in [1,n]$$

where the shares ω_i and v_i for all $i \in [0,n]$, as well as the tax rate r, are the control variables. The state variables are of course K_i^p, K_i^g and P_i for all $i \in [0,n]$, as well as Φ. We have taken the only technological progress to be public capital augmenting, although clearly other options exist.

7.3.5 Numerical Example Solved by Time Discretization

As a numerical example, we consider four regions as shown in Figure 7.1. In this example, we assume there is no technological for the sake of simplicity and numerical tractability.We employ a Cobb-Douglas production function for each region, which is expressed as

$$Y_i = F_i\left(K_i^p, \Phi K_i^g, P_i\right) = A_i (K_i^p \Phi(t) K_i^g)^{1-\alpha} (P_i)^{\alpha}$$

where A_i is a productivity parameter and α is a coefficient. We assume α is 0.5 and A_i for each region is shown in Table 7.1. We set the private and public capital decay rates of δ_p and δ_g at 0.005 and 0.01, respectively and the constant nominal rate of discount ρ at 0.05. Note that the fitness measure is

$$\zeta_i = V_i\left(K_i^g, t\right) + \Psi_i(t)$$

where $V_i\left(K_i^g, t\right)$ describes the effect of infrastructure on carrying capacity and $\Psi_i(t)$ is the natural or ambient carrying capacity which exists in the absence of infrastructure investment. To make this example simple, we assume that $V_i\left(K_i^g, t\right)$ is linearly proportional to K_i^g and $\Psi_i(t)$ is constant:

$$V_i\left(K_i^g, t\right) = v_i K_i^g$$

$$\Psi_i(t) = \psi_i$$

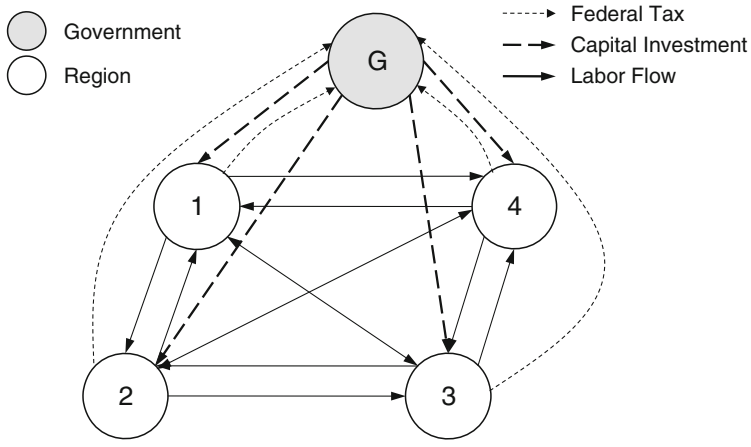


Fig. 7.1 Capital, labor, and tax flows

Table 7.1 Numerical values of parameters

Parameter	Region 1	Region 2	Region 3	Region 4
A_i	0.100	0.120	0.090	0.125
c_i	0.10	0.20	0.08	0.30
γ_i	0.01	0.02	0.03	0.02
λ_i	0.01	0.03	0.005	0.01
v_i	0.05	0.03	0.01	0.05
ψ_i	100	120	110	150
ξ_i	0.01	0.03	0.05	0.02
κ_i	0.02	0.01	0.06	0.05

The remaining numerical values of parameters used in this example are given in Table 7.1. A complete presentation of this example's mathematical formulation, associated data, algorithmic details and numerical solution may be found by following self-explanatory links found at the website http://www2.ie.psu.edu/csee/DODG/Ch7opt.pdf.

7.4 Exercises

1. Refering to the optimal control model (7.148), (7.149), (7.150), (7.151), (7.152), (7.153), (7.154), and (7.155), show the necessary conditions are also sufficient, and then determine the optimal policy $\left(r^*, \mu_1^*, \ldots, \mu_n^*\right)$.
2. Note that the dynamics of technological progress

$$\frac{d\Phi}{dt} = \tau \left[\Theta - \sigma \left(K^p, K^g, L, \Phi\right)\right]_+$$

given by expression (7.175) involve a nondifferentiable righthand side. Hence, the associated Hamiltonian will also be nondifferentiable. One technique for dealing with this circumstance is to replace (7.175) with

$$\frac{d\Phi}{dt} = \tau \left\{ [\Theta - \sigma (K^p, K^g, L, \Phi)]_+ \right\}^2$$

Discuss whether these alternative dynamics sacrifice any modeling rigor. Explain why these alternative dynamics will allow the maximum principle to be applied in the form presented in previous chapters and analyze the associated optimality conditions to obtain investment decision rules like those developed for the Harrod-Domar family of models in Section 7.2 of this chapter.

3. In Section 7.3.4, we introduce the notion of full labor force participation, so that $L_i = P_i$ for every region $i \in [1, n]$. Provide an alternative model formulation that relaxes the assumption of full labor force participation. Analyze the associated optimality conditions.
4. For the model of Section 7.3.4, introduce transportation in form of mobility of capital by assuming transportation costs a fixed percentage of the appropriate form of capital. Analyze the optimality conditions of the resulting model, assuming there is no technological progress and $\Phi = 1$ has a constant value.
5. Develop and present a discrete-time approximation of an extension of the example of Section 7.3.5 that uncludes technological progress. Solve your approximation using a commercial nonlinear program solver, such as MINOS. Comment on your solution.
6. Provide regularity conditions for the models of this chapter that assure the optimal control necessary conditions that have been invoked are also sufficient.

List of References Cited and Additional Reading

Arrow, K. J. and M. Kurz (1970). *Public Investment, the Rate of Return, and Optimal Fiscal Policy*. Baltimore: The Johns Hopkins University Press.

Bagchi, A. (1984). *Stackelberg Differential Games in Economic Models*. Berlin: Springer-Verlag.

Bazaraa, M., H. Sherali, and C. Shetty (1993). *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton: Princeton University Press.

Datta-Chaudhuri, M. (1967). Optimum allocation of investments and transportation in a two-region economy. In K. Shell (Ed.), *Essays on the Theory of Optimal Economic Growth*, pp. 129–140. MIT Press.

Domazlicky, B. (1977). A note on the inclusion of transportation in models of the regional allocation of investment. *Journal of Regional Science 17*, 235–241.

Friesz, T. and N. Kydes (2003). The dynamic telecommunications flow routing problem. *Networks and Spatial Economics 4*(1), 55–73.

Friesz, T. and J. Luque (1987). Optimal regional growth models: multiple objectives, singular controls, and sufficiency conditions. *Journal of Regional Science 27*, 201–224.

Hadley, G. and M. Kemp (1971). *Variational Methods in Economics*. North-Holland Amsterdam.

Hahn, F. and R. Matthews (1964). The theory of economic growth: a survey. *The Economic Journal 74*(296), 779–902.

Hotelling, H. (1978). A mathemathical theory of population. *Environment and Planning A 10*, 1223–1239.

Intriligator, M. (1964). Regional allocation of investment: comment. *Quarterly Journal of Economics 78*, 659–662.

Ohtsuki, Y. A. (1971). Regional allocation of public investment in an *n*-region economy. *Journal of Regional Science 11*, 225–233.

Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko (1962). *The Mathematical Theory of Optimal Processes*. New York: Interscience.

Puu, T. (1989). *Lecture Notes in Economics and Mathematical Systems*. New York: Springer-Verlag.

Puu, T. (1997). *Mathematical Location and Land Use Theory; An Introduction*. New York: Springer-Verlag.

Rahman, M. (1963a). Regional allocation of investment: An aggregative study in the theory of development programming. *The Quarterly Journal of Economics 77*(1), 26–39.

Rahman, M. (1963b). Regional allocation of investment: an aggregative study in the theory of development programming. *Quarterly Journal of Economics 77*, 26–39.

Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal 38*, 543–559.

Sakashita, N. (1967a). Regional allocation of public investment. *Papers in Regional Science 19*(1), 161–182.

Sakashita, N. (1967b). Regional allocation of public investments. *Papers, Regional Science Association 19*, 161–182.

Sethi, S. P. and G. L. Thompson (1981). *Optimal Control Theory: Applications to Management Science*. Boston: Martinus Nijhoff.

Tait, K. (1965). *Singular Problems in Optimal Control*. Ph. D. thesis, Harvard University, Cambridge, MA.

Takayama, A. (1967). Regional allocation of investment: a further analysis. *The Quarterly Journal of Economics*, 330–337.

Chapter 8
Production Planning, Oligopoly and Supply Chains

In this chapter we develop models that describe how prices, production rates and distribution activities evolve over time and influence one another for three output market structures:

1. perfect competition
2. monopoly, and
3. oligopoly.

In particular, we apply the material from previous chapters to the modeling and computation of production, distribution, and supply chain decisions made by firms operating within the three competitive environments mentioned above. Throughout this chapter our perspective is deterministic, and the dynamic games considered are open loop in nature with perfect initial information. We begin with aspatial models and move to models with explicit network path flows. We shall deal exclusively with finite terminal times and see that policies near the terminal time are of great importance to the lifetime profitability of firms. One of our goals will be to study how policies on inventory remaining at the terminal time as will as the value of such residual inventories when liquidated can influence operations throughout a firm's history.

The following is a preview of the principal topics covered in this chapter:

Section 8.1: The Aspatial Price-Taking Firm. In this section, we present a simple dynamic production planning model involving a single firm.

Section 8.2: The Aspatial Monopolistic Firm. In this section, we present a model of dynamic production planning for a monopoly.

Section 8.3: The Monopolistic Firm in a Network Economy. In this section, we extend the model of Section 8.2 to a network structure.

Section 8.4: Dynamic Oligopolistic Spatial Competition. In this section, we present a network model of dynamic oligopolistic competition, wherein firms are located at nodes and distribution of output takes place over routes that are sequences of arcs.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_8, © Springer Science+Business Media, LLC 2010

Section 8.5: Competitive Supply Chains. In this section, we present an integrated supply-production-distribution model along with an informative numerical example.

8.1 The Aspatial Price-Taking Firm

We first consider the circumstance of perfect competition. The price of the single homogeneous output of each firm is a known function of continuous time

$$\pi(t) \in L^2\left[t_0, t_f\right]$$

since we assume every firm conducts its business in a perfectly competitive market and is, as a consequence, a price-taker. Moreover, as mentioned above, the decision environment is deterministic and open loop. The time interval considered is $\left[t_0, t_f\right] \subseteq \Re^1_+$, where $t_0 \in \Re^1_+$ is the fixed initial time, $t_f \in \Re^1_{++}$ is the fixed terminal time, and of course $t_f > t_0$. There is a constant nominal rate of discounting ρ, and compounding is continuous. The firm's output rate is $q(t)$ with associated production cost $V(q)$; the rate of allocation of output to consumption is $c(t)$; and the firm's inventory is $I(t)$, a quantity of undelivered stock or a quantity of backorders according to its sign. Moreover, the controls

$$q \in L^2\left[t_0, t_f\right]$$

$$c \in L^2\left[t_0, t_f\right]$$

completely determine the state (inventory) which may be viewed as the operator

$$I(q, c) : L^2\left[t_0, t_f\right] \times L^2\left[t_0, t_f\right] \longrightarrow \mathcal{H}^1\left[t_0, t_f\right]$$

where $L^2\left[t_0, t_f\right]$ is the space of square-integrable functions and $\mathcal{H}^1\left[t_0, t_f\right]$ is a Sobolev space for the real interval $\left[t_0, t_f\right] \in \Re^1_+$. The upper bounds on output, consumption and inventory are, respectively

$$Q \in \Re^1_{++}$$

$$C \in \Re^1_{++}$$

$$K \in \Re^1_{++}$$

In addition we use the notation

$$\psi(I) : \mathcal{H}^1\left[t_0, t_f\right] \longrightarrow \mathcal{H}^1\left[t_0, t_f\right]$$

for the inventory holding/backorder cost functional.

8.1.1 Optimal Control Problem for Aspatial Perfect Competition

A consequence of the above notation and assumptions is the firm's extremal problem:

$$\max\ p_f I\left(t_f\right) e^{-\rho t_f} + \int_{t_0}^{t_f} e^{-\rho t} \left\{\pi\left(t\right) c - V\left(q\right) - \Psi\left(I\right)\right\} dt \tag{8.1}$$

subject to

$$\frac{dI}{dt} = q - c \tag{8.2}$$

$$I\left(0\right) = I_0 \tag{8.3}$$

$$0 \leq I \leq K \tag{8.4}$$

$$0 \leq q \leq Q \tag{8.5}$$

$$0 \leq c \leq C \tag{8.6}$$

where p_f is the price per unit of inventory liquidated at the terminal time.Within the criterion, $p_f I\left(t_f\right) e^{-\rho t_f}$ is the present value of inventory and backorders at the terminal time. Under the integral is the instantaneous net present value of profit for the firm, consisting of the price multiplied by the amount consumed minus the variable cost of production for each instant in time. This is integrated to give the net present value of the time stream of profits. If p_f is high enough, there may be inventory held until the terminal time t_f to take advantage of a favorable liquidation price, one that might have been negotiated well in advance.

8.1.2 Numerical Example of Aspatial Perfect Competition

As an example let us assume that inventory and backorder costs are zero, while the remaining model parameters are

$$\begin{aligned} t_0 &= 0 \\ t_f &= 10 \\ \pi\left(t\right) &= 1 + 0.01t \\ V\left(q\right) &= \frac{1}{2}q^2 \\ \Psi\left(I\right) &= 0 \\ p_f &= 1 \\ \rho &= 0.05 \\ C &= Q = 10 \end{aligned}$$

Also, for simplicity, we will completely relax the constraints

$$0 \leq I \leq K$$

Consequently we face the optimal control problem

$$\begin{aligned} \max\ & I\,(10)\, e^{-(0.05)(10)} + \int_0^{10} e^{-(0.05)t} \left[(1 + 0.01t)\, c - \frac{1}{2} q^2 \right] dt \\ & \frac{dI}{dt} = q - c \\ & I\,(0) = 100 \\ & 0 \leq q \leq 10 \\ & 0 \leq c \leq 10 \end{aligned}$$

We see immediately that c will have a bang-bang solution because it appears linearly in the problem formulation; there may or may not be a singular consumption control c. The Hamiltonian is

$$\begin{aligned} H\,(c, q, \lambda) &= e^{-\rho t} \{\pi\,(t)\, c - V\,(q)\} + \lambda\,(q - c) \\ &= e^{-(0.05)t} \left\{ (1 + 0.01t)\, c - \frac{1}{2} q^2 \right\} + \lambda\,(q - c) \end{aligned}$$

So we have

$$\begin{aligned} & \frac{d\lambda}{dt} = -\frac{\partial H}{\partial I} = 0 \\ & \quad \Rightarrow \lambda^* = \lambda\,(10) = \frac{\partial \left[p_f\, I\,(10)\, e^{-(0.05)(10)} \right]}{\partial I\,(10)} = e^{-(0.5)} \end{aligned}$$

If $0 < q < 10$ we have

$$\frac{\partial H}{\partial q} = e^{-(0.05)t}\,(-q) + \lambda = 0$$

Hence

$$q^*\,(t) = \left[\lambda e^{(0.05)t} \right]_0^{10}$$

Note that for $t \in [0, 10]$

$$10 > e^{(0.05)t - (0.5)} > 0$$

so we know

$$q^*(t) = e^{(0.05)t - (0.5)}$$

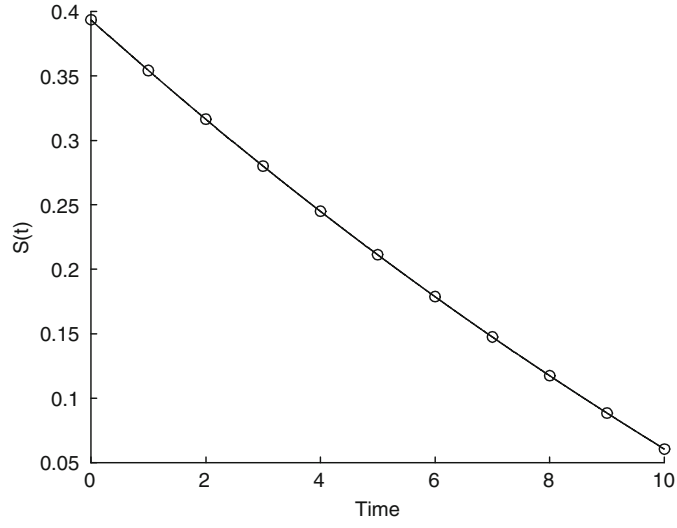
We also know that c obeys

$$c = \begin{cases} C & \text{if } S > 0 \\ 0 & \text{if } S < 0 \\ c_s & \text{if } S = 0 \end{cases}$$

where S is the switching function given by

$$\begin{aligned} S = \frac{\partial H}{\partial c} &= e^{-(0.05)t}\,(1 + 0.01t) - \lambda \\ &= e^{-(0.05)t}\,(1 + 0.01t) - e^{-(0.5)} \end{aligned}$$

whose plot is



We can see that S is positive on the interval $[0, 10)$, so $c^* = C = 10$ since the set of times for which the switching function vanishes on $[0, 10]$ is a singleton and has measure zero. That is, there is no singular control c_s.

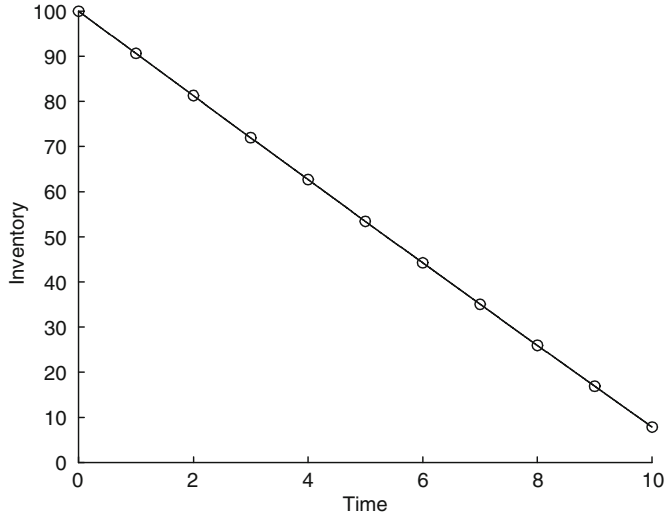
It is now a simple matter to determine the optimal production schedule and the optimal inventory history. In particular, inventory is determined from the initial-value problem

$$\begin{aligned} \frac{dI}{dt} &= e^{(0.05)t-(0.5)} - 10 \\ I\,(0) &= 100 \end{aligned}$$

whose solution is

$$I\,(t) = 12.131 \exp[0.05t] - 10.0t + 87.869$$

The plot of inventory versus time is



which makes clear that $I(10) > 0$ and the liquidation value of residual inventory make it desirable to hold some stock until the terminal time.

8.1.3 The Aspatial Price Taking Firm with a Terminal Constraint on Inventory

An alternative treatment of inventory/backordering is to stipulate that inventory is zero at the terminal time but may change sign as often as needed prior to the terminal time to support profit maximization by the firm. In that case the abstract model is

$$\max \eta I\left(t_f\right) e^{-\rho t_f} + \int_{t_0}^{t_f} e^{-\rho t} \left\{\pi\left(t\right) c - V\left(q\right) - \Psi\left(I\right)\right\} dt$$

subject to

$$\begin{aligned} \frac{dI}{dt} &= q - c \\ I\left(0\right) &= I_0 \\ 0 &\leq q \leq Q \\ 0 &\leq c \leq C \end{aligned}$$

where η is the dual variable associated with the terminal time constraint

$$I\left(t_f\right) = 0$$

that assures no unfulfilled backorders. Consequently, we could drop the residual value term $\eta I\left(t_f\right) e^{-\rho t_f}$ from our formulation. However, it is instructive to retain the residual term and see what happens.

Let us consider the same data used for the example of Section 8.1 but now with some important additions:

$$\Psi(I) = \frac{1}{2} I^2 \tag{8.7}$$

$$I\left(t_f\right) = 0 \tag{8.8}$$

Consequently, we face the optimal control problem

$$\max\ \eta I\,(10)\, e^{-(0.05)(10)} + \int_0^{10} e^{-(0.05)t}\left[(1+0.01t)\,c - \frac{1}{2}q^2 - \frac{1}{2}I^2\right] dt$$

$$\begin{aligned} \frac{dI}{dt} &= q - c \\ I\,(0) &= 100 \\ 0 &\le q \le 10 \\ 0 &\le c \le 10 \end{aligned}$$

where η is a dual variable for the terminal time constraint (8.8) on inventory. The Hamiltonian is

$$\begin{aligned} H(c,q,\lambda) &= e^{-\rho t}\left\{\pi(t)\,c - V(q)\right\} + \lambda(q-c) \\ &= e^{-(0.05)t}\left[(1+0.01t)\,c - \frac{1}{2}q^2 - \frac{1}{2}I^2\right] + \lambda(q-c) \end{aligned}$$

So we have

$$\frac{d\lambda}{dt} = (-1)\frac{\partial H}{\partial I} = e^{-(0.05)t} I$$

$$\lambda\,(10) = \frac{\partial\left[\eta I\,(10)\, e^{-(0.05)(10)}\right]}{\partial I\,(10)} = \eta e^{-(0.05)(10)}$$

As before if $0 < q < 10$ we have

$$\frac{\partial H}{\partial q} = e^{-(0.05)t}(-q) + \lambda = 0$$

Hence the production control law is

$$q(t) = \left[\lambda e^{(0.05)t}\right]_0^{10}$$

We again know that c is a potentially bang-bang and possibly singular control:

$$c = \begin{cases} 10 & \text{if } S > 0 \\ 0 & \text{if } S < 0 \\ c_s & \text{if } S = 0 \end{cases} \tag{8.9}$$

where S is the switching function given by

$$S = \frac{\partial H}{\partial c} = e^{-(0.05)t}(1 + 0.01t) - \lambda = 0,$$

and it is not presently known whether there is a singular control c_s. Let us assume

$$S > 0 \tag{8.10}$$

$$q \leq 10 \tag{8.11}$$

for all $t \in [0, 10]$. Hence the control laws are

$$c = 10 \tag{8.12}$$

$$q = \max\left(0, \lambda e^{(0.05)t}\right) \tag{8.13}$$

so that the state and adjoint equations form the two-point boundary-value problem:

$$\begin{aligned} \frac{dI}{dt} &= \max(0, \lambda e^{(0.05)t}) - 10 \\ \frac{d\lambda}{dt} &= e^{-(0.05)t} I \\ I(0) &= 100 \\ \lambda(10) &= \eta e^{-(0.05)(10)} \end{aligned}$$

Note that one of the boundary conditions involves the unknown dual variable η. As such the problem would appear to be very challenging. However, we know the terminal value of the state variable; that is, $I(10) = 0$. As a consequence we seek to find the value η of the final-value problem

$$\begin{aligned} \frac{dI}{dt} &= \max\left(0, \lambda e^{(0.05)t}\right) - 10 \\ \frac{d\lambda}{dt} &= e^{-(0.05)t} I \\ I(10) &= 0 \\ \lambda(10) &= \eta e^{-(0.05)(10)} \end{aligned}$$

using a *reverse shooting method* wherein we adjust the value η to force the initial state value $I\ (0) = 100$. Let us select $\eta = 0$ so that

$$\lambda\ (10) = \eta e^{-(0.05)(10)} = 0$$

and assume that $\lambda(t) < 0$ for $t \in [0, 10)$; in which case inventory must obey

$$\begin{aligned} \frac{dI}{dt} &= -10 \\ I\ (10 - \epsilon) &= 0 \end{aligned}$$

as $\epsilon \longrightarrow 0$, with the apparent closed-form solution

$$I\ (t) = 100 - 10t$$

which also satisfies $I(0) = 100$. It follows that the adjoint dynamics have the simplified form

$$\begin{aligned} \frac{d\lambda}{dt} &= e^{-(0.05)t}\ (100 - 10t) \\ \lambda\ (10) &= 0 \end{aligned}$$

so that

$$\lambda\ (t) = \exp(-.05t)\ (200t + 2000) - 2426.1 \tag{8.14}$$

whose plot, the reader may easily verify, agrees with our assumption that the adjoint is negative for $t \in [0, 10)$. Note further that the transversality condition is satisfied:

$$\begin{aligned} \lambda\ (10) &= \exp(-.05(10))\ (200(10) + 2000) - 2426.1 \\ &= 4000 \exp(-.5) - 2426.1 \\ &= 2426.1 - 2426.1 = 0.0 \end{aligned}$$

We also note that the nonpositivity of the adjoint variable requires that

$$S = e^{-(0.05)t}\ (1 + 0.01t) - \lambda = e^{-(0.05)t}\ (1 + 0.01t) + |\lambda| > 0$$

in keeping with (8.9), (8.10), and (8.12). So that the consumption rate is

$$c = 10$$

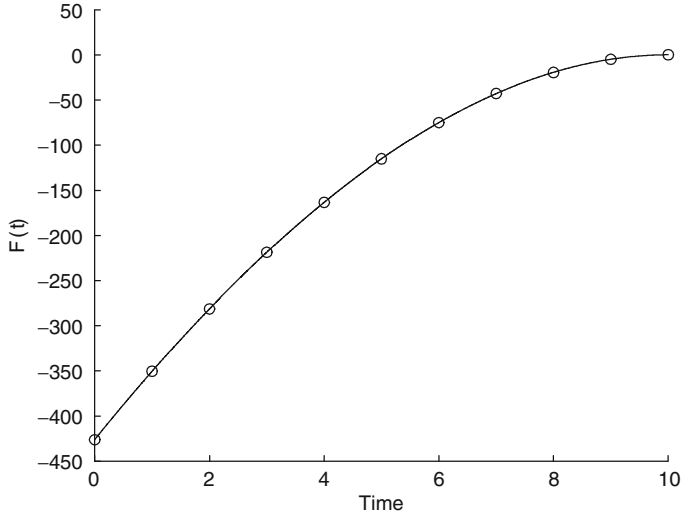
for all time $t \in [0, 10]$ and there are no singular controls. Furthermore, the optimal production policy is

$$q\ (t) = \max(0, \lambda e^{(0.05)t}) = \max(0, F\ (t))$$

where

$$F(t) \equiv \exp(.05t)\left[\exp(-.05t)\left(200t + 2000\right) - 2426.1\right]$$

Note that $F\left(t\right) \leq 0$ as is evident from the plot:



Consequently,

$$q\left(t\right) = 0$$

for all time $t \in [0, 10]$, which is understood to mean that the firm sells from its inventories and does not produce any finished products during the time interval of our analysis.

8.2 The Aspatial Monopolistic Firm

We now turn our attention to aspatial monopoly in a dynamic setting. The firm of interest is now without competition and exploits the demand curve for its product in order to maximize profit. We employ the notion of an inverse demand curve; that is,

$$\pi = \theta\left(c, t\right)$$

denotes the price consumers pay for the firm's product as a function of consumption rate c and time t. The explicit dependence on time arises from changes in consumer preferences over time and/or seasonal effects. Consequently, we consider the following model:

$$\max\ J = \eta I\left(t_f\right) e^{-\rho t_f} + \int_{t_0}^{t_f} e^{-\rho t}\left\{\theta\left(c, t\right) c - V\left(q\right) - \Psi\left(I\right)\right\} dt$$

subject to

$$\begin{aligned} \frac{dI}{dt} &= q - c \\ I(0) &= I_0 \\ I\left(t_f\right) &= 0 \\ 0 \leq q &\leq Q \\ 0 \leq c &\leq C \end{aligned}$$

where once again η is the dual variable for the terminal condition $I\left(t_f\right) = 0$.

8.2.1 Necessary Conditions for the Aspatial Monopoly

The relevant Hamiltonian is

$$H = e^{-\rho t}\left[\theta(c,t)\, c - V(q) - \Psi(I)\right] + \lambda(q - c) \tag{8.15}$$

Provided the Hamiltonian is convex in its controls, the minimum principal requires

$$\begin{aligned} c^* &= \left[\arg\left(\frac{\partial H}{\partial c} = 0\right)\right]_0^C \\ q^* &= \left[\arg\left(\frac{\partial H}{\partial q} = 0\right)\right]_0^Q \end{aligned}$$

The adjoint equation and transversality condition are

$$\frac{d\lambda}{dt} = (-1)\frac{\partial H}{\partial I} = e^{-\rho t}\frac{d\Psi(I)}{dI} \tag{8.16}$$

$$\lambda\left(t_f\right) = \frac{\partial \eta I\left(t_f\right) e^{-\rho t_f}}{\partial I(t_f)} = \eta e^{-\rho t_f} \tag{8.17}$$

Assuming none of the control constraints are binding for the optimal trajectory, we have

$$\frac{\partial H}{\partial c} = e^{-\rho t}\frac{\partial}{\partial c}\left[\theta(c,t)\, c\right] - \lambda = 0 \tag{8.18}$$

$$\frac{\partial H}{\partial q} = -e^{-\rho t}\frac{\partial V(q)}{\partial q} + \lambda = 0 \tag{8.19}$$

which tells us

$$\frac{\partial}{\partial c}\left[\theta(c,t)\, c\right] = \frac{\partial V(q)}{\partial q} = \lambda e^{\rho t} \tag{8.20}$$

along the optimal trajectory. This last expression is a statement that marginal revenue with respect to consumption equals marginal variable cost with respect to output. Futhermore, the adjoint variable is the present value of marginal variable cost with respect to output:

$$\lambda = e^{-\rho t} \frac{\partial V(q)}{\partial q}$$

8.2.2 Numerical Example

We will consider an example based on the data

$$\begin{aligned} t_0 &= 0 \\ t_f &= 10 \\ \theta(c,t) &= (11-c)\exp\left(\frac{1}{10}t\right) \\ V(q) &= \frac{1}{2}q^2 \\ \Psi(I) &= \frac{1}{2}I^2 \\ I(t_f) &= 0 \\ \rho &= 0.05 \\ C &= Q = 10 \end{aligned}$$

which is identical to our example of Section 8.1.3 except that the firm, now a monopoly, may exploit the market's inverse demand function $\theta(c,t)$ which describes price as falling when consumption increases as well as generally drifting upward over time (as a result of wage growth or other economic forces). Thus, we face the optimal control problem

$$\max\ \eta I(10)e^{-(0.05)(10)} + \int_0^{10} e^{-(0.05)t}\left\{(11-c)\exp\left(\frac{1}{10}t\right)c - \frac{1}{2}q^2 - \frac{1}{2}I^2\right\}dt$$

subject to

$$\begin{aligned} \frac{dI}{dt} &= q - c \\ I(0) &= I_0 \\ 0 &\le q \le 10 \\ 0 &\le c \le 10 \end{aligned}$$

The associated Hamiltonian is

$$H(c,q,\lambda) = e^{-0.05t}\left\{\left(11c - c^2\right)e^{\frac{1}{10}t} - \frac{1}{2}q^2 - \frac{1}{2}I^2\right\} + \lambda(q-c)$$

Note that the problem is no longer linear in c and, as such, c will not be bang-bang or singular. Moreover if $0 < c < 10$ then

$$\frac{\partial H}{\partial c} = e^{-0.05t}(11 - 2c)\, e^{\frac{1}{10}t} - \lambda = 0$$

so the consumption control law is

$$c = \left[-\frac{1}{2} \frac{-11e^{-0.05t} e^{\frac{1}{10}t} + \lambda}{e^{-0.05t} e^{\frac{1}{10}t}} \right]_0^{10}$$

$$= \min\left(10, \max\left(0, -\frac{1}{2} \frac{-11e^{-0.05t} e^{\frac{1}{10}t} + \lambda}{e^{-0.05t} e^{\frac{1}{10}t}}\right)\right)$$

Also if $0 < q < 10$ then

$$\frac{\partial H}{\partial q} = e^{-0.05t}(-q) + \lambda = 0$$

so the production control law is

$$q = \left[\lambda e^{0.05t}\right]_0^{10} = \min(10, \max(0, \lambda e^{0.05t}))$$

The adjoint equation and terminal time condition are:

$$\frac{d\lambda}{dt} = (-1)\frac{\partial H}{\partial I} = e^{-(0.05)t} I$$

$$\lambda(10) = \frac{\partial \eta I(10) e^{-(0.05)(10)}}{\partial I(10)} = \eta e^{-(0.05)(10)}$$

Consequently, the relevant two-point boundary-value problem is

$$\frac{dI}{dt} = \min\left(10, \max\left(0, \lambda e^{0.05t}\right)\right) - \min\left(10, \max\left(0, -\frac{1}{2} \frac{-11e^{-0.05t} e^{\frac{1}{10}t} + \lambda}{e^{-0.05t} e^{\frac{1}{10}t}}\right)\right)$$

$$\frac{d\lambda}{dt} = e^{-(0.05)t} I$$

$$I(0) = 100$$

$$\lambda(10) = \eta e^{-(0.05)(10)}$$

Note that one of the boundary conditions involves the unknown dual variable η. As such the problem can be restated as a final-value problem and solved using a reverse shooting method. Let us select the multiplier η so that

$$\lambda(10) = \eta e^{-(0.05)(10)} = -100$$

giving the system

$$\frac{dI}{dt} = \min\left(10, \max\left(0, \lambda e^{0.05t}\right)\right) - \min\left(10, \max\left(0, -\frac{1}{2}\frac{-11e^{-0.05t}e^{\frac{1}{10}^t} + \lambda}{e^{-0.05t}e^{\frac{1}{10}^t}}\right)\right)$$

$$\frac{d\lambda}{dt} = e^{-(0.05)t} I$$

$$I(10) = 0$$

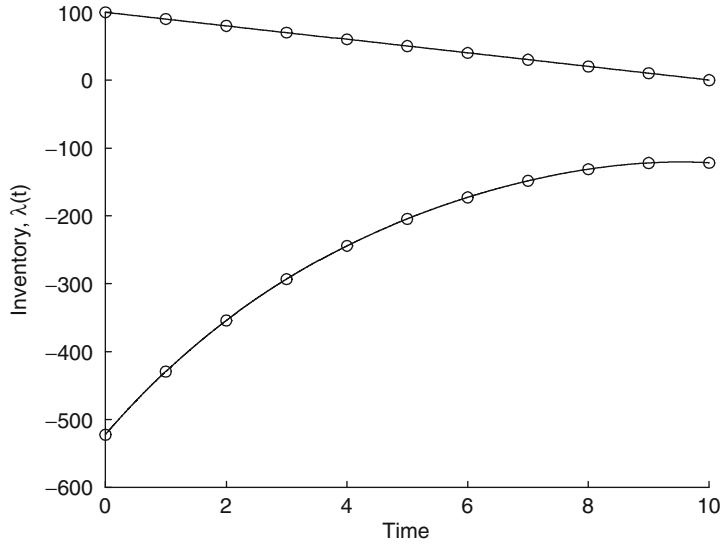
$$\lambda(10) = -100$$

whose exact solution is

$$I = -10.0t + 100.0$$

$$\lambda = e^{-0.05t}(200.0t + 2000.0) - 100 - 4000\exp(-.5)$$

This solution has the following graphical expression:



where the straight line is inventory $I(t)$ and the other curve is the adjoint variable $\lambda(t)$. Clearly inventory is never negative, and $I(10) = 0$ so that there are no

unfulfilled backorders. Note also that $\lambda(t) < 0$, so again

$$q = \left[\lambda e^{0.05t}\right]_0^{10} = 0$$

for all $t \in [0, 10]$.

8.3 The Monopolistic Firm in a Network Economy

Now we consider the firm of interest to be located at one of the nodes of a distribution network for which flows over paths generate flows over arcs controlled by shipping agents who set freight tariffs for the firm's output at the origin-destination (OD) pair level. Although some of the notation needed is identical or very similar to notation already introduced in this chapter, we provide a self-contained treatment of notation within this section to avoid any confusion. In particular, time is again denoted by the scalar $t \in \Re^1_+$, initial time by $t_0 \in \Re^1_+$, final time by $t_f \in \Re^1_{++}$, with $t_0 < t_f$ so that $t \in \left[t_0, t_f\right] \subset \Re^1_+$. There are three sets important to articulating a model of production and distribution on a network; these are as follow: $\mathcal{A}$ for directed arcs, $\mathcal{N}$ for nodes, and $\mathcal{W}$ for origin-destination (OD) pairs. Subsets of these sets are formed as is meaningful by using the subscripts i for a specific node and ij for a specific OD pair (i, j).

The firm controls production output rates expressed as a vector q, allocations of output to meet demand expressed as a vector c, and shipping patterns expressed as a vector s. Inventories I are a vector of state variables determined by the controls. That is,

$$\begin{aligned} c &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|} \\ q &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|} \\ s &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{W}|} \\ I(c, q, s) : \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|} &\times \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|} \times \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{W}|} \\ &\longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{N}|} \end{aligned}$$

where again $L^2\left[t_0, t_f\right]$ is the space of square-integrable functions and $\mathcal{H}^1\left[t_0, t_f\right]$ is a Sobolev space for the real interval $\left[t_0, t_f\right] \in \Re^1_+$.

8.3.1 The Network Firm's Extremal Problem

The firm has the objective of maximizing net profit expressed as revenue less cost and taking the form of an operator acting on production rates, shipment patterns, and allocations of output to meet demands. For simplicity we imagine that the firm

has operations at every node and that every node is a perfectly competitive market for the firm's output. That is, the firm's net profit is

$$J(c,q,s) = \int_{t_0}^{t_f} e^{-\rho t} \left\{ \sum_{i\in\mathcal{N}} \pi_i (c_i, t)\, c_i - \sum_{i\in\mathcal{N}} V_i(q,t) - \sum_{(i,j)\in\mathcal{W}} r_{ij}(t)\, s_{ij} - \sum_{i\in\mathcal{N}} \psi_i(I_i,t) \right\} dt \tag{8.21}$$

where $\rho \in \Re^1_{++}$ is a constant nominal rate of discount, $r_{ij} \in L^2_{++}\left[t_0, t_f\right]$ is the exogenous freight rate (tariff) charged per unit of flow s_{ij} for OD pair $(i,j) \in \mathcal{W}$,

$$\psi_i : \mathcal{H}^1\left[t_0, t_f\right] \times \Re^1_+ \longrightarrow \Re^1_+$$

is the firm's inventory cost at node i, and I_i is the inventory/backorder at node i. In (8.21), c_i is the allocation of the firm's output to consumption at node i. Our formulation is in terms of flows and is based on the inverse demand functions[1]

$$\pi_i (c_i, t) : L^2\left[t_0, t_f\right] \times \Re^1_+ \longrightarrow \Re^1_+$$

Furthermore, q_i is the firm's output at node $i \in \mathcal{N}$. Also

$$V_i(q,t) : \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|} \times \Re^1_+ \longrightarrow \Re^1_+$$

is the variable cost of production for the firm at node $i \in \mathcal{N}$. Note that $J(c,q,s)$ is a functional that is completely determined by the controls. The first term of the functional $J(c,q,s)$ in expression (8.21) is the firm's revenue; the second term is the firm's cost of production; the third term is the firm's shipping costs; and the last term is the firms's inventory or holding cost.

We also impose the terminal time inventory constraints

$$I_i\left(t_f\right) = K_i \quad \forall\, i \in \mathcal{N} \tag{8.22}$$

where the $K_i \in \Re^1_{++}$ are exogenous. All consumption, production, and shipping control variables are nonnegative and bounded from above. That is,

$$C \geq c \geq 0 \tag{8.23}$$

$$Q \geq q \geq 0 \tag{8.24}$$

$$S \geq s \geq 0 \tag{8.25}$$

[1] The assumption of demand separability is easily relaxed.

where

$$C \in \Re_{++}^{|\mathcal{F}|}$$
$$Q \in \Re_{++}^{|\mathcal{F}|}$$
$$S \in \Re_{++}^{|\mathcal{W}|}$$

are known constant vectors. Constraints (8.23), (8.24), and (8.25) are recognized as pure control constraints, while (8.22) are terminal conditions for the state variables. Naturally,

$$\Omega = \{(c, q, s) : (8.23), (8.24), (8.25)\} \tag{8.26}$$

is the set of feasible controls.

The inventory dynamics, expressing simple flow conservation, obey

$$\frac{dI_i}{dt} = q_i + \sum_{(j,i)\in\mathcal{W}} s_{ji} - \sum_{(i,j)\in\mathcal{W}} s_{ij} - c_i \quad \forall\, i \in \mathcal{N} \tag{8.27}$$

$$I_i(t_0) = I_i^0 \quad \forall\, i \in \mathcal{N} \tag{8.28}$$

where $I_i^0 \in \Re_{++}^1$ is exogenous. Consequently, the vector of inventories may be viewed as the operator

$$I(c,q,s) = \arg\left\{ \frac{dI_i}{dt} = q_i + \sum_{(j,i)\in\mathcal{W}} s_{ji} - \sum_{(i,j)\in\mathcal{W}} s_{ij} - c_i, \right.$$
$$\left. I_i(t_0) = I_i^0,\ \ I_i(t_f) = K_i \ \ \forall i \in \mathcal{N} \right\}$$

where we implicitly assume that the dynamics have solutions for all feasible controls.

We are nearly ready to give a succinct statement of the extremal problem faced by the firm carrying out its production and distribution activities to meet demands on a network. The firm solves an optimal control problem to determine its production q, allocation of output to meet demand c, and shipping pattern s by maximizing its profit functional $J(c,q,s)$ subject to inventory dynamics expressed as flow balance equations and upper and lower bounds for the controls (c,q,s). With the preceding development, we note that the firm's problem is: compute c, q and s (thereby finding I) in order to solve

$$\begin{aligned} &\max\ J(c,q,s) \\ &\text{s.t.}\quad (c,q,s) \in \Omega \end{aligned} \tag{8.29}$$

where Ω is as defined in (8.26).

8.3.2 Discrete-Time Approximation

We note that (8.29) can be solved in a number of ways, although direct appeal to the necessary conditions is unlikely to be successful for general networks due to the large number of variables. Furthermore, since there are no time shifts and the state dynamics are linear in the formulation proposed above, time discretization and finite-dimensional mathematical programming is especially appealing. To this end we construct the following discrete-time approximation of the criterion of (8.29), where t now denotes a discrete time period:

$$J(c,q,s) = \left\{ \sum_{k=0}^{N} e^{-\rho t_k} \sum_{i \in \mathcal{N}} \pi_i \left[c_i\left(t_k\right), t_k \right] c_i\left(t_k\right) - \sum_{i \in \mathcal{N}} V_i\left(\left[q\left(t_k\right), t_k\right]\right. \right.$$
$$\left. \left. - \sum_{(i,j) \in \mathcal{W}} r_{ij}\left(t_k\right) s_{ij}\left(t_k\right) - \sum_{i \in \mathcal{N}} \psi_i\left(\left[I_i\left(t_k\right), t_k\right]\right) \right\} \tag{8.30}$$

The discrete-time approximation of the associated dynamics is

$$I_i\left(t_k\right) - I_i\left(t_{k-1}\right) = q_i\left(t_k\right) + \sum_{(j,i) \in \mathcal{W}} s_{ji}\left(t_k\right) - \sum_{(i,j) \in \mathcal{W}} s_{ij}\left(t_k\right) - c_i\left(t_k\right)$$
$$\forall\, i \in \mathcal{N}, \forall k \in [1, N] \tag{8.31}$$

where we define

$$t = t_0 + j\Delta t$$

and N is the number of discretizations, defined by

$$N = \frac{t_f - t_0}{\Delta t}$$

Furthermore

$$I_i\left(t_0\right) = I_i^0 \qquad \forall\, i \in \mathcal{N} \tag{8.32}$$

and

$$\begin{aligned} c\left(t_k\right) &= \left(c_i\left(t_k\right) : i \in \mathcal{N}\right) \\ q\left(t_k\right) &= \left(q_i\left(t_k\right) : i \in \mathcal{N}\right) \\ s\left(t_k\right) &= \left(s_{ij}\left(t_k\right) : (i,j) \in \mathcal{W}\right) \\ c &= \left(c\left(t_k\right) : k \in [1, N]\right) \\ q &= \left(q\left(t_k\right) : k \in [1, N]\right) \\ s &= \left(s\left(t_k\right) : k \in [1, N]\right) \end{aligned}$$

so that we may write

$$C \geq c \geq 0 \tag{8.33}$$
$$Q \geq q \geq 0 \tag{8.34}$$
$$S \geq s \geq 0 \tag{8.35}$$

8.3.3 Numerical Example

Let us consider a network of five arcs and four nodes for which the single firm of interest has activities located at each node $i = 1, 2, 3, 4$. Consumption of the firm's output potentially occurs at every node; this consumption may be of local output or of imported output as the network topology permits. Figure 8.1 illustrates the network. The time interval of interest is $[0, 10]$; that is $t_0 = 0$ and $t_f = 10$. Before time discretization, there are 13 controls and 4 state variables associated with this example; these are listed in Table 8.1.

At time $t_0 = 0$, the initial inventory at each node is

$$I_1(0) = 5 \tag{8.36}$$
$$I_2(0) = 3 \tag{8.37}$$
$$I_3(0) = 2$$
$$I_4(0) = 0 \tag{8.38}$$

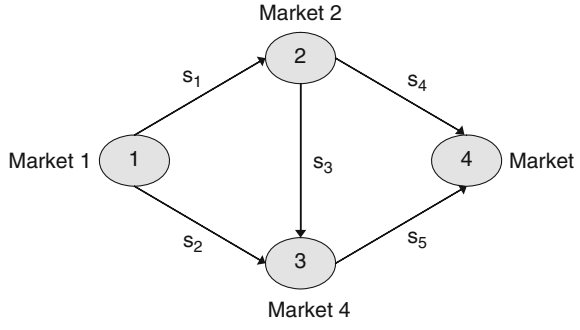


Fig. 8.1 Network of five arcs, four nodes

Table 8.1 Controls and states for example

Controls			States
c_1	q_1	s_1	I_1
c_2	q_2	s_2	I_2
c_3	q_3	s_3	I_3
c_4	q_4	s_4	I_4
-	-	s_5	-

In addition, we impose the condition that no backordering is allowed by any firm at any node at the terminal time $t_f = 10$. That is,

$$I_i(10) = 0 \text{ for } i = 1, 2, 3, 4 \tag{8.39}$$

The inventory dynamics are the following flow balance equations:

$$\begin{aligned}
\frac{dI_1}{dt} &= q_1 - s_1 - s_2 - c_1 \\
\frac{dI_2}{dt} &= q_2 + s_1 - s_3 - s_4 - c_2 \\
\frac{dI_3}{dt} &= q_3 + s_2 + s_3 - s_5 - c_3 \\
\frac{dI_4}{dt} &= q_4 + s_4 + s_5 - c_4
\end{aligned}$$

We assume the inverse demands at each node take the following form:

$$\left.\begin{aligned}
\pi_1(c_1, t) &= 4.0(11 - c_1)\exp\left(\frac{t}{40}\right) \qquad \pi_2(c_2, t) = 3.0(11 - c_2)\exp\left(\frac{t}{30}\right) \\
\pi_3(c_3, t) &= 3.5(11 - c_3)\exp\left(\frac{t}{35}\right) \qquad \pi_4(c_4, t) = 2.5(11 - c_4)\exp\left(\frac{t}{25}\right)
\end{aligned}\right\} \tag{8.40}$$

The production cost functions for each node have the form

$$\left.\begin{aligned}
V_1(q_1, t) &= 1.00(q_1)^2 \qquad V_3(q_3, t) = 0.65(q_3)^2 \\
V_2(q_2, t) &= 0.35(q_2)^2 \qquad V_4(q_4, 5) = 2.00(q_4)^2
\end{aligned}\right\} \tag{8.41}$$

We assume the holding costs are

$$\left.\begin{aligned}
\psi_1(I_1, t) &= 0.5\,(I_1)^2 \qquad \psi_3(I_3, t) = 1.5\,(I_3)^2 \\
\psi_2(I_2, t) &= 5.0\,(I_2)^2 \qquad \psi_4(I_4, t) = 2.0\,(I_4)^2
\end{aligned}\right\} \tag{8.42}$$

We assume that the freight rates for each arc are the following constants:

$$\left.\begin{aligned}
r_1(s_1, t) &= 5 \qquad r_3(s_3, t) = 3 \qquad r_5(s_5, t) = 4 \\
r_2(s_2, t) &= 2 \qquad r_4(s_4, t) = 2
\end{aligned}\right\} \tag{8.43}$$

We also impose the following bounds on control variables:

$$C = \begin{pmatrix} 5 \\ 10 \\ 10 \\ 5 \end{pmatrix} \quad Q = \begin{pmatrix} 5 \\ 2 \\ 5 \\ 5 \end{pmatrix} \quad S = \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \end{pmatrix}$$

The individual firms' profit functions are found by substituting (8.40), (8.41), (8.42), and (8.43) into (8.21), then discretizing to obtain a form like (8.30).

8.3.4 Solution by Discrete-Time Approximation

We solve the functional mathematical program corresponding to the details presented in Section 8.3.3 using a discrete-time approximation with $N = 10$ equal time steps. In our calculations we allowed GAMS/MINOS to solve the specific instance of (8.30) through (8.35) corresponding to the data we have given. The solution time for this example is approximately 2 cpu seconds on a Pentium® 4 single-processor computer. The results are presented in Figure 8.2.

8.3.5 Solution by Continuous-Time Gradient Projection

We next solve the example problem presented in Section 8.3.3 by the continuous-time gradient projection method. We calculate 40 values of the gradients and then construct a 6-th order polynomial approximation of each as a smooth function of time. The algorithm is implemented in MATLAB and the solution time for the example presented is approximately 10 cpu seconds on a Pentium® 4 single-processor computer. The results are shown in Figure 8.3. The gradient projection algorithm is articulated below in terms of the state and control vectors specific to the example problem:

Gradient Projection Algorithm

Step 0. Initialization. Set $k = 0$ and pick $c_i^0(t)$, $q_i^0(t)$, and $s_i^0(t)$ for $i = 1, 2, 3, 4$ where time t is now continuous.

Step 1. Find state trajectory. Using current controls, solve the state initial-value problem

$$\frac{dI_1}{dt} = q_1^k - s_1^k - s_2^k - c_1^k \qquad I_1(0) = 5$$
$$\frac{dI_2}{dt} = q_2^k + s_1^k - s_3^k - s_4^k - c_2^k \quad I_2(0) = 3$$

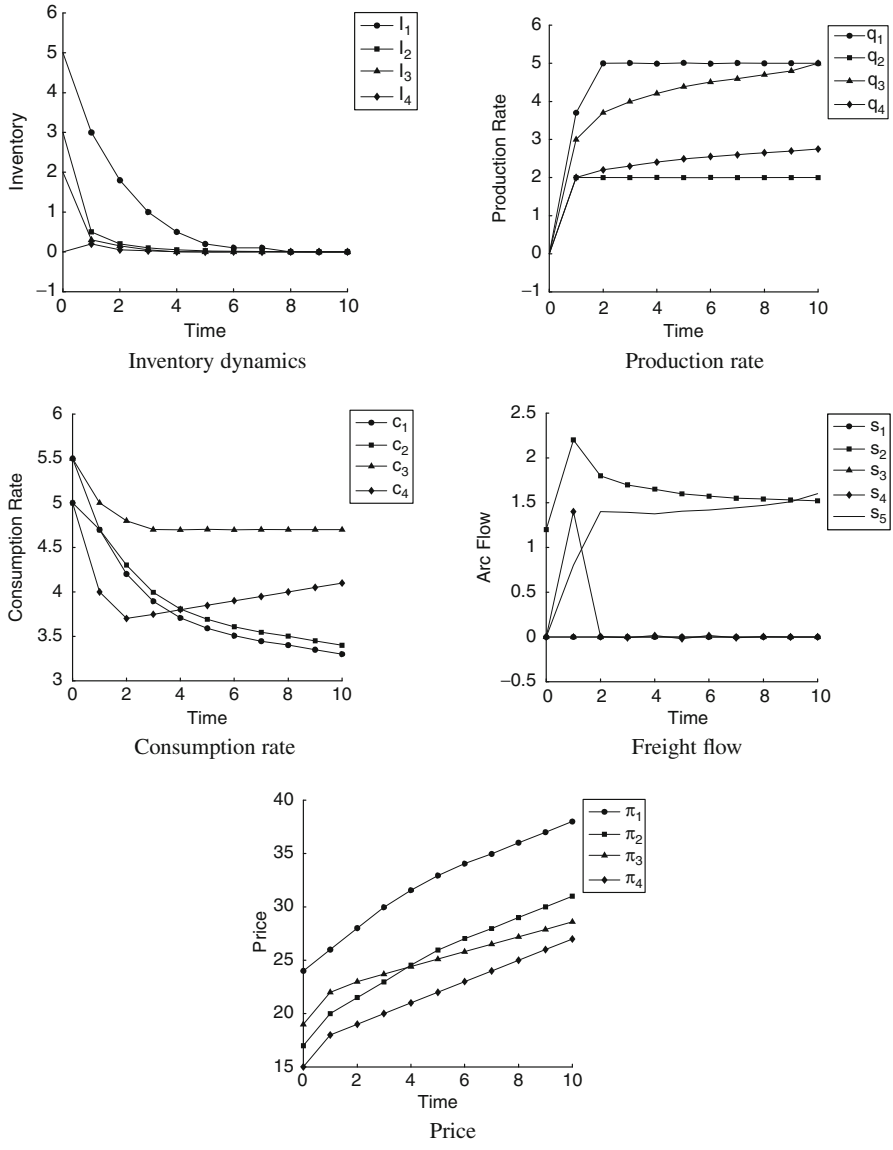


Fig. 8.2 Solution by discrete-time approximation

$$\frac{dI_3}{dt} = q_3^k + s_2^k + s_3^k - s_5^k - c_3^k \quad I_3(0) = 2$$
$$\frac{dI_4}{dt} = q_4^k + s_4^k + s_5^k - c_4^k \quad I_4(0) = 0$$

and call the solution $I_1^k(t)$, $I_2^k(t)$, $I_3^k(t)$, and $I_4^k(t)$.

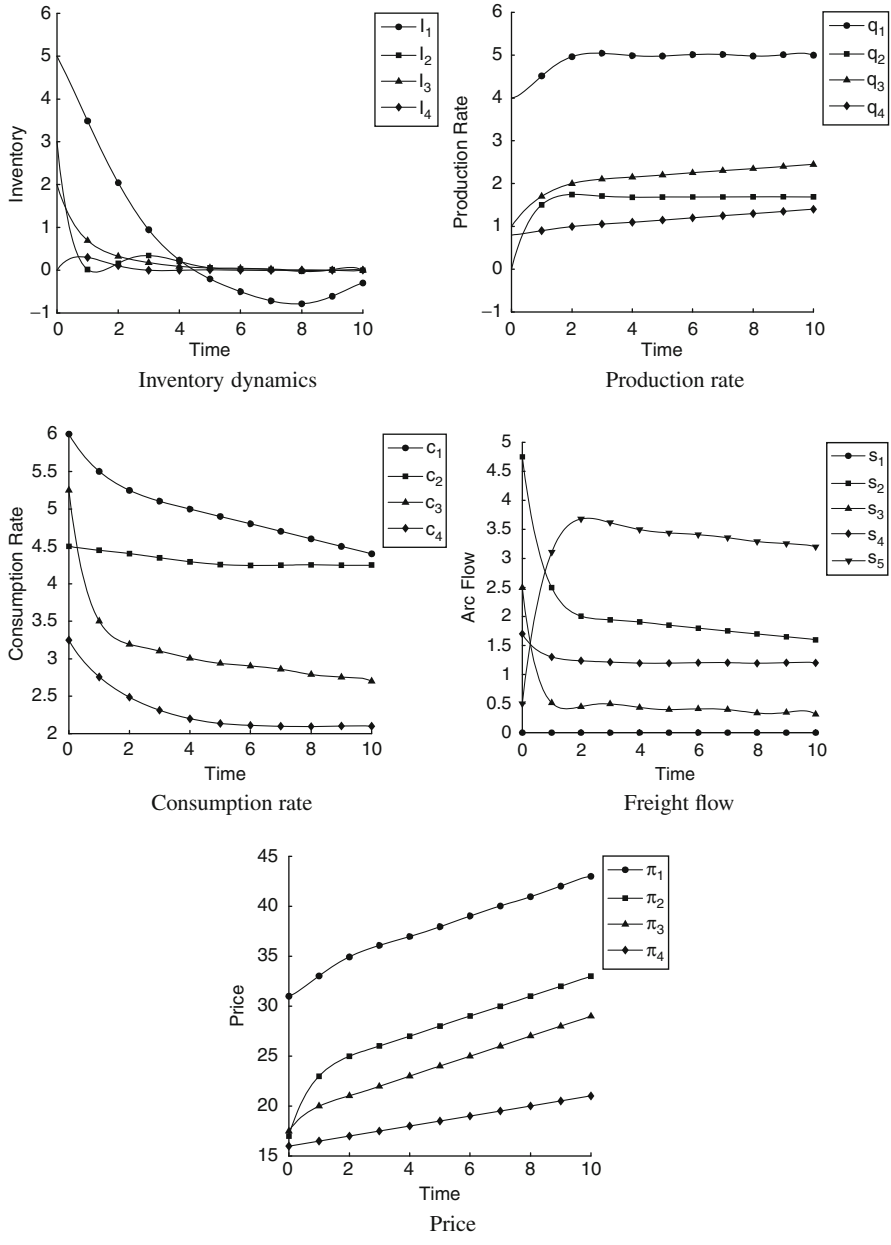


Fig. 8.3 Solution by continuous time gradient projection method

Step 2. Find adjoint trajectory. Using current controls and states, solve the adjoint final-value problem

$$(-1)\frac{d\lambda_1}{dt} = \exp(-\rho t)(I_1^k) \qquad \lambda_1(10) = \eta I_1^k(10)$$

$$(-1)\frac{d\lambda_2}{dt} = \exp(-\rho t)(10I_2^k) \qquad \lambda_2(10) = \eta I_2^k(10)$$

$$(-1)\frac{d\lambda_3}{dt} = \exp(-\rho t)(3I_3^k) \qquad \lambda_3(10) = \eta I_3^k(10)$$

$$(-1)\frac{d\lambda_4}{dt} = \exp(-\rho t)(4I_4^k) \qquad \lambda_4(10) = \eta I_4^k(10)$$

picking the dual variable η to enforce zero inventory at the terminal time; call the solution $\lambda_1^k(t), \lambda_2^k(t), \lambda_3^k(t)$, and $\lambda_4^k(t)$.

Step 3. Find gradient. Using current controls, states, and adjoints, calculate

$$\nabla_u J(u^k) = \nabla_u H\left(I^k, u^k, \lambda^k, t\right) = \nabla_u f_0(I^k, u^k, t) + \nabla_u\left[\lambda^T f\left(I^k, u^k, t\right)\right]$$

where

$$u^k = \begin{pmatrix} c^k \\ q^k \\ s^k \end{pmatrix}$$

and $H(I, u, \lambda, t)$ is the relevant Hamiltonian for this problem.

Step 4. Update and apply stopping test. For a suitably small step size θ_k, update according to

$$u^{k+1} = P_U\left[u^k - \theta_k \nabla J\left(u^k\right)\right]$$

If an appropriate stopping test is satisfied, declare

$$u^*(t) \approx u^{k+1}(t)$$

Otherwise, set $k = k + 1$ and go to Step 1.

8.4 Dynamic Oligopolistic Spatial Competition

Recently Friesz et al. (2006) showed that it is possible to model differential Nash equilibria among producers whose facilities and final demand markets are fixed at distinct nodes of a distribution network and connected by paths involving chains of arcs of that network. Their model, which takes the form of a differential variational inequality, is presented in this section.

8.4.1 Some Background and Notation

The oligopolistic firms of interest, embedded in a network economy, are in oligopolistic competition according to dynamics that describe the trajectories of inventories/backorders and correspond to flow conservation for each firm at each node of the network of interest. The oligopolistic firms, acting as shippers, compete as price takers in the market for physical distribution services which is perfectly competitive. Perfect competition in shipping arises because numerous shipping companies serve numerous customers due to the involvement of shippers in the numerous output markets of the network economy. The time scale we consider is neither short nor long, but rather of sufficient length to allow output and shipping pattern adjustments although not long enough for firms to relocate or enter or leave the network economy.

We employ much of the notation introduced in previous sections of this chapter. Because there are some key differences between the dynamic ologopolistic competition to now be studied and problems explored previously in this chapter, we choose to give an exhaustive list of the notation to be employed below, even though that will involve a bit of duplication. In particular, we again let continuous time be denoted by the scalar $t \in \Re^1_+$, initial time by $t_0 \in \Re^1_+$, and final time by $t_f \in \Re^1_{++}$, with $t_0 < t_f$ so that $t \in [t_0, t_f] \subset \Re^1_+$. There are several sets important to articulating a model of oligopolistic competition on a network; these are as follow: $\mathcal{F}$ for firms, $\mathcal{A}$ for directed arcs, $\mathcal{N}$ for nodes and $\mathcal{W}$ for origin-destination (OD) pairs. Subsets of these sets are formed as is meaningful by using the subscripts f for a specific firm, i for a specific node, and ij for a specific OD pair (i, j).

Each firm $f \in \mathcal{F}$ controls production (output) rates q^f, allocation of output to meet demand c^f and shipping pattern s^f. Inventories I^f are state variables determined by the controls. In particular, concatenations of the firm-specific vectors c^f, q^f, and s^f give the following:

$$
\begin{aligned}
c &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \\
q &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \\
s &\in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{W}|\times|\mathcal{F}|}
\end{aligned}
$$

Furthermore, the state operator, once again, will be

$$
I\,(c,q,s) : \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \times \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|} \times \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{W}|\times|\mathcal{F}|} \longrightarrow \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{N}|\times|\mathcal{F}|}
$$

where $L^2\left[t_0, t_f\right]$ is the space of square-integrable functions and $\mathcal{H}^1\left[t_0, t_f\right]$ is a Sobolev space for the real interval $\left[t_0, t_f\right] \in \Re^1_+$.

8.4.2 The Firm's Objective and Constraints

Each firm has the objective of maximizing net profit expressed as revenue less cost and taking the form of an operator acting on allocations of output to meet demands, production rates, and shipment patterns. For each firm $f \in \mathcal{F}$, net profit is

$$\begin{aligned}\Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f}) &= e^{-\rho t_f} Z_f\left[I(t_f), t_f\right] \\ &+ \int_{t_0}^{t_f} e^{-\rho t} \left\{ \sum_{i \in \mathcal{N}} \pi_i \left(\sum_{g \in \mathcal{F}} c_i^g, t \right) c_i^f \right. \\ &- \sum_{i \in \mathcal{N}_f} V_i^f(q^f, t) - \sum_{(i,j) \in \mathcal{W}_f} r_{ij}(t)\, s_{ij}^f \\ &- \sum_{i \in \mathcal{N}} \psi_i^f(I_i^f, t) dt \end{aligned} \tag{8.44}$$

where $\rho \in \Re^1_{++}$ is a constant nominal rate of discount, $r_{ij} \in \Re^1_{++}$ is the freight rate (tariff) charged per unit of flow s_{ij} for OD pair $(i, j) \in \mathcal{W}_f$, ψ_i^f is firm f's inventory cost at node i, and I_i^f is the inventory/backorder of firm f at node i. In (8.44), c_i^f is the allocation of the output of firm $f \in \mathcal{F}$ at node $i \in \mathcal{N}$ to consumption at that node. Also

$$Z_f\left[I^f(t_f), t_f\right]$$

is the liquidation value of inventory remaining at the terminal time, where

$$I^f = \left(I_i^f : i \in \mathcal{N}_f\right)$$

Our formulation is in terms of flows so we employ the inverse demand functions $\pi_i(c_i, t)$ where

$$c_i = \sum_{g \in \mathcal{F}} c_i^g$$

is the total allocation of output to consumption for node i. Furthermore, q_i^f is the output of firm $f \in \mathcal{F}$ at node $i \in \mathcal{N}$. Also $V_i^f(q, t)$ is the variable cost of production for firm $f \in \mathcal{F}$ at node $i \in \mathcal{N}$. Note that $\theta_f(c^f, q^f, s^f; c^{-f}, q^{-f})$ is a functional that is completely determined by the controls c^f, q^f and s^f when non-own allocations to consumption and non-own production rates

$$\begin{aligned} c^{-f} &\equiv (c^{f'} : f' \neq f) \\ q^{-f} &\equiv (q^{f'} : f' \neq f) \end{aligned}$$

are taken as exogenous data by firm f. The first term of the objective functional $\theta_f(c^f, q^f, s^f; c^{-f}, q^{-f})$ in expression (8.44) is the firm's revenue; the second

term is the firm's cost of production; the third term is the firm's shipping costs; and the last term is the firm's inventory or holding cost.

We also impose the terminal time inventory constraints

$$I_i^f\left(t_f\right) = \tilde{K}_i^f \quad \forall\, f \in \mathcal{F}, i \in \mathcal{N}_f \tag{8.45}$$

where the $\tilde{K}_i^f \in \Re^1_{++}$ are exogenous. All consumption, production, and shipping variables are nonnegative and bounded from above; that is,

$$C^f \geq c^f \geq 0 \tag{8.46}$$

$$Q^f \geq q^f \geq 0 \tag{8.47}$$

$$S^f \geq s^f \geq 0 \tag{8.48}$$

where

$$C^f \in \Re^{|\mathcal{F}|}_{++}$$

$$Q^f \in \Re^{|\mathcal{F}|}_{++}$$

$$S^f \in \Re^{|\mathcal{W}_f|}_{++}$$

Constraints (8.46), (8.47), and (8.48) are recognized as pure control constraints, while (8.45) are terminal conditions for the state space variables. Naturally

$$\Omega_f = \left\{\left(c^f, q^f, s^f\right) : (8.46), (8.47), (8.48)\right\}$$

is the set of feasible controls.

Firm f solves an optimal control problem to determine its production q^f, allocation of production to meet demand c^f, and shipping pattern s^f – thereby also determining inventory I^f via dynamics we articulate momentarily – by maximizing its profit functional $\Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f})$ subject to inventory dynamics expressed as flow balance equations and pertinent production and inventory constraints. The inventory dynamics for firm $f \in \mathcal{F}$, expressing simple flow conservation, obey

$$\frac{dI_i^f}{dt} = q_i^f + \sum_{(j,i)\in\mathcal{W}} s_{ji}^f - \sum_{(i,j)\in\mathcal{W}} s_{ij}^f - c_i^f \quad \forall\, i \in \mathcal{N}_f \tag{8.49}$$

$$I_i^f\,(t_0) = K_i^f \quad \forall\, i \in \mathcal{N}_f \tag{8.50}$$

$$I_i^f\left(t_f\right) = \tilde{K}_i^f \quad \forall\, i \in \mathcal{N}_f \tag{8.51}$$

where $K_i^f \in \Re^1_{++}$ and $\tilde{K}_i^f \in \Re^1_+$ are exogenous. Note that the transportation time for the flow of finished goods is not captured explicitly in the inventory dynamics, however it is accounted for implicitly in the freight rate (tariff) charged per unit of

flow. Further, in addition to the terminal time inventory (state) constraints (8.51), the model is general enough to handle inventory constraints over the entire planning horizon $\left[t_0, t_f\right]$. For instance, nonnegativity of the inventory (state) variables could be imposed to restrict firms from taking backorders. In light of the above notions and definitions, we may write

$$I\left(c,q,s\right) = \arg\left\{\frac{dI_i^f}{dt} = q_i^f + \sum_{(j,i)\in\mathcal{W}} s_{ji}^f - \sum_{(i,j)\in\mathcal{W}} s_{ij}^f - c_i^f, \right.$$
$$\left. I_i^f\left(t_0\right) = K_i^f, \;\; I_i^f\left(t_f\right) = \tilde{K}_i^f \;\; \forall f \in \mathcal{F} \;\; i \in \mathcal{N}_f \right\}$$

where we implicitly assume that the dynamics have solutions for all feasible controls.

With the preceding development, we note that firm f's problem is: with the c^{-f} and q^{-f} as exogenous inputs, compute c^f, q^f and s^f (thereby finding I^f) in order to solve the following extremal problem:

$$\left.\begin{array}{l} \max\ \Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f}) \\ \text{s.t.} \quad \left(c^f, q^f, s^f\right) \in \Omega_f \end{array}\right\} \forall f \in \mathcal{F} \tag{8.52}$$

where

$$\Omega_f = \left\{\left(c^f, q^f, s^f\right) : (8.45), (8.46), (8.47), (8.48) \text{ hold}\right\}$$

also for all $f \in \mathcal{F}$. That is, each firm is a Nash agent that knows and employs the current instantaneous values of the decision variables of other firms to make its own noncooperative decisions. As such, (8.52) is a differential Nash game.

8.4.3 The DVI Formulation

We assume this game is regular in the sense of the following definition:

Definition 8.1. *The dynamic oligopolistic network competition problem introduced above will be considered regular if: (1) the state operator $I\left(c,q,s\right)$ exists and is unique, while each of its components is continuous and G-differentiable; (2) the inverse demand, production cost and inventory cost functions are continuously differentiable with respect to controls and states; and (3) for each $f \in \mathcal{F}$, the composite terminal cost function*

$$Z_f\left[I^f\left(t_f\right), t_f\right] + \sum_{i\in\mathcal{N}_f} \gamma_i^f \left[\tilde{K}_i^f - I_i^f\left(t_f\right)\right]$$

is continuously differentiable with respect to $I_i^f\left(t_f\right)$ for all $i \in \mathcal{N}_f$.

In the above definition, each γ_i^f is a constant dual variable that prices out the terminal constraint on inventory.

We further note that (8.52) is an optimal control problem with fixed terminal time. Its Hamiltonian is

$$H_f\left(c^f, q^f, s^f, I^f, \alpha^f, \beta^f, \lambda^f; c^{-f}; q^{-f}; t\right)$$

$$\equiv \Phi_f\left(c^f, q^f, s^f, I^f; c^{-f}, q^{-f}; t\right) + \Psi_f\left(c^f, q^f, s^f, I^f, \alpha^f, \beta^f, \lambda^f\right)$$

where

$$\begin{aligned}\Phi_f\left(c^f, q^f, s^f, I^f; c^{-f}, q^{-f}; t\right) = e^{-\rho t}\Bigg\{&\sum_{i\in\mathcal{N}_f}\pi_i\left(\sum_{g\in\mathcal{F}} c_i^g, t\right)c_i^f \\ &- \sum_{i\in\mathcal{N}_f} V_i^f(q,t) - \sum_{(i,j)\in\mathcal{W}_f} r_{ij}(t)\, s_{ij}^f \\ &- \sum_{i\in\mathcal{N}_f}\psi_i^f(I_i^f, t)\Bigg\}\end{aligned} \tag{8.53}$$

and

$$\begin{aligned}&\Psi_f\left(c^f, q^f, s^f, I^f, \alpha^f, \beta^f, \lambda^f\right) \\ &\quad = \sum_{i\in\mathcal{N}_f}\lambda_i^f\left(q_i^f + \sum_{(j,i)\in\mathcal{W}} s_{ji}^f - \sum_{(i,j)\in\mathcal{W}} s_{ij}^f - c_i^f\right)\end{aligned} \tag{8.54}$$

where $\alpha_i^f \in \Re_+^1$ and $\beta_i^f \in \Re_+^1$ are dual variables for the inventory-bounding constraints (8.45) while $\alpha^f \in \Re^{|\mathcal{N}_f|}$ and $\beta^f \in \Re^{|\mathcal{N}_f|}$; also $\lambda_i^f \in \Re_+^1$ is the adjoint variable for the dynamics of firm f at node i while $\lambda^f \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{N}|}$. Clearly Φ_f is the instantaneous profit. To interpret Ψ_f we need to understand the relevant dynamic shadow benefits and shadow costs of this model. To that end, recall that, along an optimal trajectory, the adjoint variables obey

$$\lambda_i^f = \frac{\partial J_f}{\partial I_i^f}$$

Consequently,

$$\Psi_f = \sum_{i\in\mathcal{N}_f}\frac{\partial J_f}{\partial I_i^f}\frac{dI_i^f}{dt}$$

which is recognized as the shadow value of dynamic benefits arising from current inventory held; it can be either a cost or a benefit, depending on its sign.

Due to regularity, the maximum principle takes the form of requiring that the nonlinear program

$$\max H_f \quad \text{s.t.} \quad \left(C^f, Q^f, S^f\right) \geq \left(c^f, q^f, s^f\right) \geq 0$$

be solved by every firm $f \in \mathcal{F}$ for every instant of time $t \in \left[t_0, t_f\right]$. Consequently, since the feasible set is convex, the finite-dimensional variational inequality principle from the necessary conditions requires any optimal solution to satisfy

$$\frac{\partial H_f^*}{\partial c_i^f}\left(c_i^f - c_i^{f*}\right) \leq 0 \tag{8.55}$$

$$\frac{\partial H_f^*}{\partial q_i^f}\left(q_i^f - q_i^{f*}\right) \leq 0 \tag{8.56}$$

$$\frac{\partial H_f^*}{\partial s_{ij}^f}\left(s_{ij}^f - s_{ij}^{f*}\right) \leq 0 \tag{8.57}$$

for every $f \in \mathcal{F}$ at every time, $t \in \left[t_0, t_f\right]$. Familiarity with variational inequalities suggests that the following variational inequality has solutions that are differential Nash equilibria for a noncooperative game in which individual firms maximize net profits in light of current information about their competitors:

$$\text{find } \left(c^{f*}, q^{f*}, s^{f*}\right) \in \Omega \text{ such that}$$

$$0 \geq \sum_{f\in\mathcal{F}} \int_{t_0}^{t_f} \left[\sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial c_i^f}\left(c_i^f - c_i^{f*}\right) + \sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial q_i^f}\left(q_i^f - q_i^{f*}\right) + \sum_{(i,j)\in\mathcal{W}_f} \frac{\partial \Phi_f^*}{\partial s_{ij}^f}\left(s_{ij}^f - s_{ij}^{f*}\right) \right] dt \quad \text{for all} \quad (c,q,s) \in \Omega \tag{8.58}$$

where

$$\Phi_f^* = \Phi_f\left(c^{f*}, q^{f*}, s^{f*}, I^{f*}; c^{-f}, q^{-f}; t\right) \tag{8.59}$$

$$\Omega = \prod_{f\in\mathcal{F}} \Omega_f \tag{8.60}$$

We note that (8.58) is a differential variational inequality expressing the differential Nash game that is our present interest. This formulation also provides guidance in devising a computational strategy, as we show in Section 8.4.4.

The issue of immediate concern is to formally demonstrate that solutions of (8.58) are differential Nash equilibria. In fact, we state and prove the following result:

Theorem 8.1. *Differential variational inequality formulation of dynamic oligopolistic network competition. Any solution of (8.58) is a solution of the dynamic oligopolistic network competition problem when regularity in the sense of Definition 8.1 holds.*

Proof. We begin by noting that (8.58) is equivalent to the following optimal control problem

$$\max G\left(c,q,s\right) = \sum_{f\in\mathcal{F}} \int_{t_0}^{t_f} \left[\sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial c_i^f} c_i^f + \sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial q_i^f} q_i^f + \sum_{(i,j)\in\mathcal{W}_f} \frac{\partial \Phi_f^*}{\partial s_{ij}^f} s_{ij}^f \right] dt$$

$$\text{s.t.} \qquad (8.45)\,,\,(8.46)\,,\,(8.47)\,,\,(8.48)\,,\,\text{and}\;\,(8.49)$$

where it is essential to recognize that $G\left(c,q,s\right)$ is a linear functional that assumes knowledge of the solution to our oligopolistic game; as such, $G\left(c,q,s\right)$ is a mathematical construct for use in analysis and has no meaning as a computational device. The augmented Hamiltonian for this artificial optimal control problem is

$$H_0 = \sum_{f\in\mathcal{F}} \left[\sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial c_i^f} c_i^f + \sum_{i\in\mathcal{N}_f} \frac{\partial \Phi_f^*}{\partial q_i^f} q_i^f + \sum_{(i,j)\in\mathcal{W}_f} \frac{\partial \Phi_f^*}{\partial s_{ij}^f} s_{ij}^f \right] + \sum_{f\in\mathcal{F}} \Psi_f$$

The associated maximum principal requires

$$\max H_0 \quad \text{s.t.} \quad \left(C^f, Q^f, S^f\right) \geq \left(c^f, q^f, s^f\right) \geq 0$$

The corresponding necessary conditions for this mathematical program are identical to (8.55) through (8.57), since

$$\frac{\partial H_0^*}{\partial c_i^f} = \frac{\partial \Phi_f^*}{\partial c_i^f} + \frac{\partial \Psi_f^*}{\partial c_i^f} = \frac{\partial H_f^*}{\partial c_i^f}$$

$$\frac{\partial H_0^*}{\partial q_i^f} = \frac{\partial \Phi_f^*}{\partial q_i^f} + \frac{\partial \Psi_f^*}{\partial q_i^f} = \frac{\partial H_f^*}{\partial q_i^f}$$

$$\frac{\partial H_0^*}{\partial s_{ij}^f} = \frac{\partial \Phi_f^*}{\partial s_{ij}^f} + \frac{\partial \Psi_f^*}{\partial s_{ij}^f} = \frac{\partial H_f^*}{\partial s_{ij}^f}$$

where

$$H_0^* = \sum_{f\in\mathcal{F}}\left[\sum_{i\in\mathcal{N}_f}\frac{\partial\Phi_f^*}{\partial c_i^f}c_i^{f*} + \sum_{i\in\mathcal{N}_f}\frac{\partial\Phi_f^*}{\partial q_i^f}q_i^{f*} + \sum_{(i,j)\in\mathcal{W}_f}\frac{\partial\Phi_f^*}{\partial s_{ij}^f}s_{ij}^{f*}\right] + \sum_{f\in\mathcal{F}}\Psi_f^*$$

and

$$\Psi_f^* = \Psi_f\left(c^{f*}, q^{f*}, s^{f*}, I^{f*}, \alpha^{f*}, \beta^{f*}, \lambda^{f*}\right)$$

■

We next note that the following existence result holds:

Theorem 8.2. *Existence of dynamic oligopolistic network equilibrium. When the variational inequality of Theorem 8.1 is regular in the sense of Definition 8.1, there exists a solution of the dynamic oligopolistic network competition problem.*

Proof. Note that each set of admissible controls Ω_f is convex and compact by the virtue of the given and the explicit lower and upper bounds of the formulation. Note also that continuity is assured by regularity. Existence is then immediate from the results of Chapter 6. ■

8.4.4 Discrete-Time Approximation

Let us define the discrete instant of time

$$t_k = t_0 + k\Delta t$$

where Δt is the time step employed, while

$$N = \frac{t_f - t_0}{\Delta t}$$

is the number of discretizations and

$$t_N = t_f$$

Then, the extremal problem (8.52) for all firms $f \in \mathcal{F}$ becomes the following:

$$\max \Phi_f(c^f, q^f, s^f; c^{-f}, q^{-f}) \approx \sum_{k=0}^{N}\tau(t_k)\, e^{-\rho t_k}\cdot\Delta \times\left\{\sum_{i\in\mathcal{N}_f}\pi_i\left(\sum_{g\in\mathcal{F}}c_i^g(t_k), t_k\right)c_i^f(t_k)\right.$$

$$- \sum_{i \in \mathcal{N}_f} V_i^f(q^f(t_k), t_k) - \sum_{(i,j) \in \mathcal{W}_f} r_{ij}(t_k)\, s_{ij}^f(t_k)$$
$$- \sum_{i \in \mathcal{N}_f} \psi_i^f(I_i^f(t_k), t_k) \Bigg\}$$

subject to

$$I_i^f(t_k) = I_i^f(t_{k-1}) + \Delta \cdot \left[q_i^f(t_k) + \sum_{(j,i) \in \mathcal{W}} s_{ji}^f(t_k) - \sum_{(i,j) \in \mathcal{W}} s_{ij}^f(t_k) - c_i^f(t_k) \right]$$
$$\forall k = 1, \ldots, N \text{ and } i \in \mathcal{N}_f$$

$$\begin{array}{ll} I_i^f(t_0) = K_i^f & \forall i \in \mathcal{N}_f \\ I_i^f(t_f) = \tilde{K}_i^f & \forall i \in \mathcal{N}_f \\ 0 \le c^f(t_k) \le C^f & \forall k \in [1, N] \\ 0 \le q^f(t_k) \le Q^f & \forall k \in [1, N] \\ 0 \le h^f(t_k) \le H^f & \forall k \in [1, N] \end{array}$$

where the vectors c^f, q^f, and h^f have the obvious definitions; moreover, $\tau(t)$ is presently the coefficient which arises from a trapezoidal approximation of the present value integral; that is

$$\tau(t) = \begin{cases} 0.5 & \text{if} \quad t = t_0 \\ 0.5 & \text{if} \quad t = t_f \\ 1 & \text{if} \quad t_0 < t < t_f \end{cases}$$

One advantage of time discretization is that we can now completely eliminate state variables (inventories) from the problem by noting that

$$I_i^f(t_{k+1}) = K_i^f + \Delta \cdot \sum_{k=0}^{t} \left[q_i^f(t_k) + \sum_{(j,i) \in \mathcal{W}} s_{ji}^f(t_k) - \sum_{(i,j) \in \mathcal{W}} s_{ij}^f(t_k) - c_i^f(t_k) \right] \tag{8.61}$$

$$I_i^f(t_f) = \tilde{K}_i^f \tag{8.62}$$

for $t = 0, \ldots, N-1$ and all $i \in \mathcal{N}_f$. As a consequence one obtains a finite-dimensional variational inequality involving only upper and lower bound constraints on the remaining control variables. This finite-dimensional variational inequality may be solved by conventional algorithms developed for such problems or a finite-dimensional nonlinear complementarity formulation may be created and used in combination with a successive linearization scheme and a linear complementarity solver.

8.4.5 A Comment About Path Variables

It should be noted that one may introduce path flows in the above formulation by re-expressing the state dynamics as

$$\frac{dI_i^f}{dt} = q_i^f + \sum_{j \in \mathcal{N}_f} \sum_{p \in \mathcal{P}_{ji}} h_p^f - \sum_{j \in \mathcal{N}_f} \sum_{p \in \mathcal{P}_{ij}} h_p^f - c_i^f$$

for every firm $f \in \mathcal{F}$ and node $i \in \mathcal{N}_f$, where $\mathcal{P}_{ji}$ is the set of paths from node $j \in \mathcal{N}_f$ to node $i \in \mathcal{N}_f$; furthermore, h_p is the flow on path $p \in \mathcal{P}_{ji}$. There are corresponding, but quite obvious, changes in the firm's objective function and the upper and lower bound constraints on its controls. We omit a complete statement of such details for the sake of brevity.

8.4.6 Numerical Example

Let us consider a network of five arcs, four nodes and four firms, where a single firm f is located at each node $i = 1, 2, 3, 4$. Consumption of each firm's output potentially occurs at every node; this consumption may be of local or of imported output as the network topology permits. Figure 8.4 illustrates the network. The time interval of interest is $[0, 20]$; that is

$$t_0 = 0$$
$$t_f = 20$$

In this example, firm 1 has an economic presence at all nodes, firm 2 at nodes 2, 3, and 4, firm 3 at nodes 3 and 4, and finally firm 4 at node 4 only. Therefore

$$\mathcal{F} = \{1, 2, 3, 4\}$$

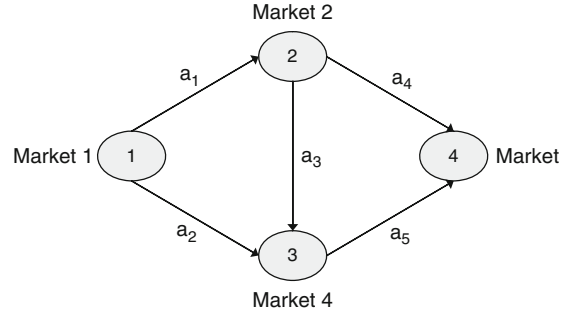


Fig. 8.4 Network of five arcs, four nodes, and four firms

and

$$\mathcal{N}_1 = \{1,2,3,4\} \quad \mathcal{N}_2 = \{2,3,4\} \quad \mathcal{N}_3 = \{3,4\} \quad \mathcal{N}_4 = \{4\}$$

Before time discretization there are 29 controls and 10 state variables associated with this example; these are enumerated in the following table:

Firm	controls by node or path										states			
1	c_1^1	c_2^1	c_3^1	c_4^1							I_1^1	I_2^1	I_3^1	I_4^1
2		c_2^2	c_3^2	c_4^2								I_2^2	I_3^2	I_4^2
3			c_3^3	c_4^3									I_3^3	I_4^3
4				c_4^4										I_4^4
all	q_1^1	q_2^2	q_3^3	q_4^4										
1	h_1^1	h_2^1	h_3^1	h_4^1	h_5^1	h_6^1	h_7^1	h_8^1	h_9^1	h_{10}^1				
2							h_7^2	h_8^2	h_9^2	h_{10}^2				
3										h_{10}^3				

At time $t_0 = 0$, every firm has an inventory of 100 units at their respective locations. That is,

$$I_i^f(0) = 100 \text{ for } f \in \mathcal{F} \text{ and } i \in \mathcal{N}_f$$

In addition, we impose the condition that no backordering is allowed by any firm at any node at the terminal time $t_f = 20$. That is

$$I_i^f(20) = 0 \text{ for } f \in \mathcal{F} \text{ and } i \in \mathcal{N}_f \tag{8.63}$$

The inventory dynamics are the flow balance equations:

$$\begin{aligned}
\frac{dI_1^1}{dt} &= q_1^1 - h_1^1 - h_2^1 - h_3^1 - h_4^1 - h_5^1 - h_6^1 - c_1^1 \\
\frac{dI_2^1}{dt} &= h_1^1 - h_7^1 - h_8^1 - h_9^1 - c_2^1 \\
\frac{dI_3^1}{dt} &= h_2^1 + h_3^1 + h_7^1 - h_{10}^1 - c_3^1 \\
&\vdots \\
\frac{dI_4^4}{dt} &= q_4^4 - c_4^4
\end{aligned} \tag{8.64}$$

which we only partially enumerate in the interest of saving space. We assume the inverse demands at each node i take the following form:

$$\pi_i(c_i, t) = \alpha_i - \beta_i (c_i)^m \tag{8.65}$$

where $m \in \Re^1_{++}$ is a constant. Also $\alpha_i \in \Re^1_{++}$ and $\beta_i \in \Re^1_{++}$ for all i are constants. The production cost functions for each firm f have the form

$$V_i^i = \frac{1}{2}\rho_i^i \left(q_i^i\right)^2 + \frac{1}{3}\sigma_i^i \left(q_i^i\right)^3 \text{ for all } i = 1, \ldots, 4 \tag{8.66}$$

where ρ_i^f and $\sigma_i^f \in \Re^1_{++}$ are also constants for all allowed i and f. In (8.66), we consider nonconvex production cost functions in order to capture both increasing and decreasing economies of scale for different production rate regimes. We assume the holding costs are quadratic and of the form

$$\psi_i^f = \frac{1}{2}\eta_i^f \left(I_i^f\right)^2 \text{ for } f \in \mathcal{F} \text{ and } i \in \mathcal{N}_f \tag{8.67}$$

where $\eta_j^i \in \Re^1_{++}$ are constants, again for allowed i and f. The relationships between arc and path variables are summarized in thee following table.

Path	Arc sequence
p_1	a_1
p_2	a_2
p_3	a_1, a_3
p_4	a_1, a_4
p_5	a_1, a_3, a_5
p_6	a_2, a_5
p_7	a_3
p_8	a_4
p_9	a_3, a_5
p_{10}	a_5

Furthermore, the relevant arc-path incidence matrix is

$$\Delta_p = \left(\delta_{ap}\right) = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

The associated path costs are

$$R = \Delta^T r \tag{8.68}$$

where

$$r = \left(r_{a_i} : i = 1, 2, 3, 4, 5\right)$$

and the

$$r_{a_i} = A_i + B_{a_i}(f_i)^n \quad i = 1, 2, 3, 4, 5$$

are unit freight rates for individual arcs and the $A_i \in \Re^1_{++}$ and $B_i \in \Re^1_{++}$ are known constants. We impose the following vectors of bounds on control variables:

$$C^f = Q^f = H^f = 75$$

Each firm's instantaneous profit function is found by substituting (8.65), (8.66), (8.67), and (8.68) into (8.53), where $\rho \in \Re^1_{++}$ is again the fixed nominal interest rate. A discrete-time approximation of the corresponding differential variational inequality was created using $N = 21$ equal time steps. The resulting finite-dimensional variational inequality was restated as a nonlinear complementarity problem and solved using GAMS with the PATH solver. The numerical values of the model's parameters are presented in the following table:

Parameter	Value	Parameter	Value	Parameter	Value
ρ	0.05	A_1	2	A_2	2
A_3	2	A_4	2	A_5	2
B_1	0.9	B_2	0.9	B_3	0.9
B_4	0.9	B_5	0.9	α_1	2000
β_1	12	α_2	2200	β_2	16
α_3	2400	β_3	14	α_4	2500
β_4	18	ρ_1^1	0.3	ρ_2^2, ρ_4^4	0.1
ρ_3^3	0.2	$\sigma_i^i, i = 1, \ldots, 4$	1	$\eta_2^1, \eta_4^1, \eta_4^3, \eta_4^4$	1
η_1^1	4	$\eta_3^1, \eta_2^2, \eta_3^2$	2	η_4^2, η_3^3	3
t_0	0	t_f	20	N	20
Δ	1	n	1	m	1

8.4.7 Interpretation of Numerical Results

Inventory trajectories are plotted against time in Figure 8.8, which shows that most firms adopt a policy of backordering at most nodes; this backordering behavior is represented by a negative inventory level. Only firms 1 and 2 have nodes that do not place backorders; these nodes correspond to the nodes where firms 1 and 2 produce their goods. The production rates for the four firms are plotted in Figure 8.5.

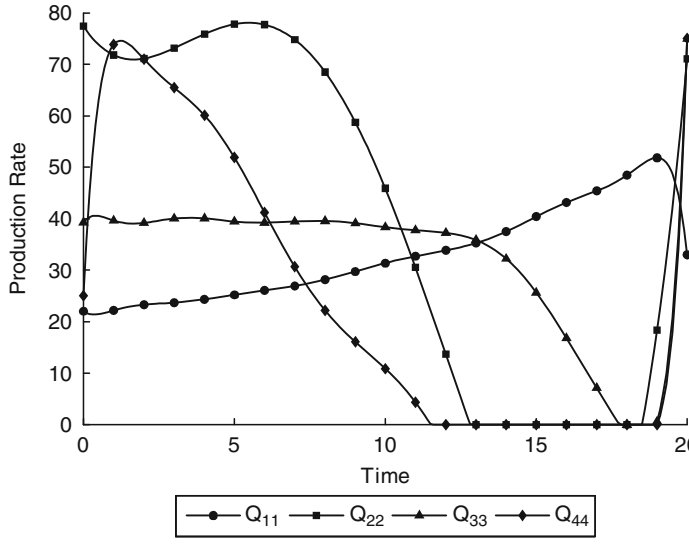


Fig. 8.5 Production rates of the firms during the planning horizon [0, 20]

Each firm seems to follow a different production plan; firm 2 operates at its full capacity for the first 10 time units, abruptly halts production and then returns to full production for the last time period to meet the final inventory constraints, whereas firm 1 slowly increases production until near the end of the planning horizon where production begins to decline. Prices of finished goods in four spatially separated markets are plotted against time in Figure 8.6. Figure 8.9 presents consumption trajectories (allocations of output to meet demands) in different markets over time. Figure 8.10 presents path flow trajectories, while Figure 8.7 illustrates the aggregated arc flows of all firms. There is relatively little transport of goods until the terminal time nears; then goods are moved between nodes to satisfy the terminal inventory constraints at each node. In Figure 8.11 we compare the net present of cumulative production, inventory holding, and transportation costs incurred by the 4 firms.The present values of profits for each firm are:

$$\begin{array}{ll} \text{Firm 1} & -\$185,592 < 0 \\ \text{Firm 2} & -\$926,070 < 0 \\ \text{Firm 3} & +\$248,179 > 0 \\ \text{Firm 4} & -\$314,978 < 0 \end{array}$$

It is evident from the above that the only firm to realize positive profits is firm 3; all other firms experience losses.

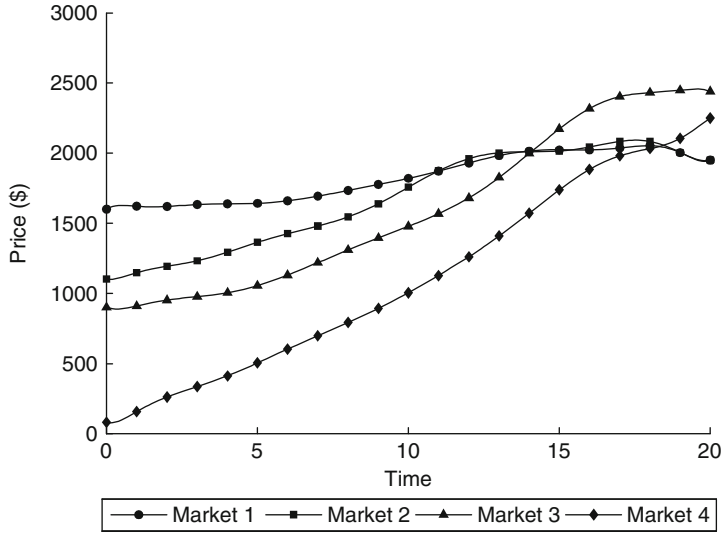


Fig. 8.6 Market price

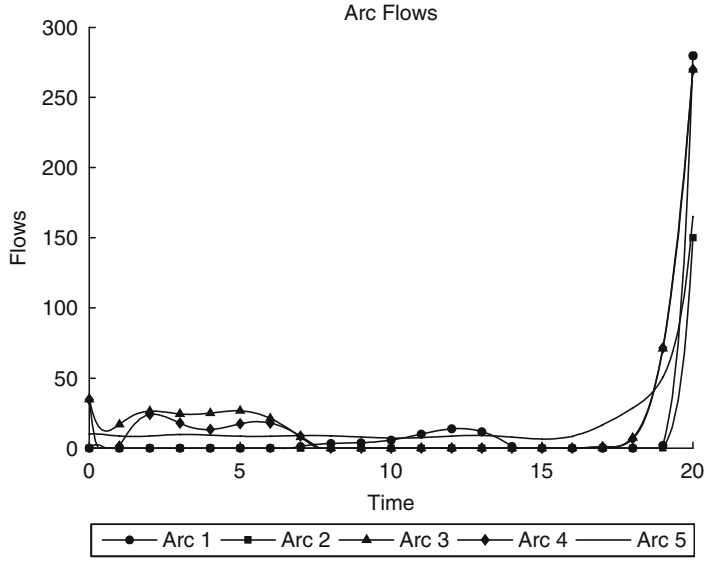


Fig. 8.7 Path flows (grouped by firms)

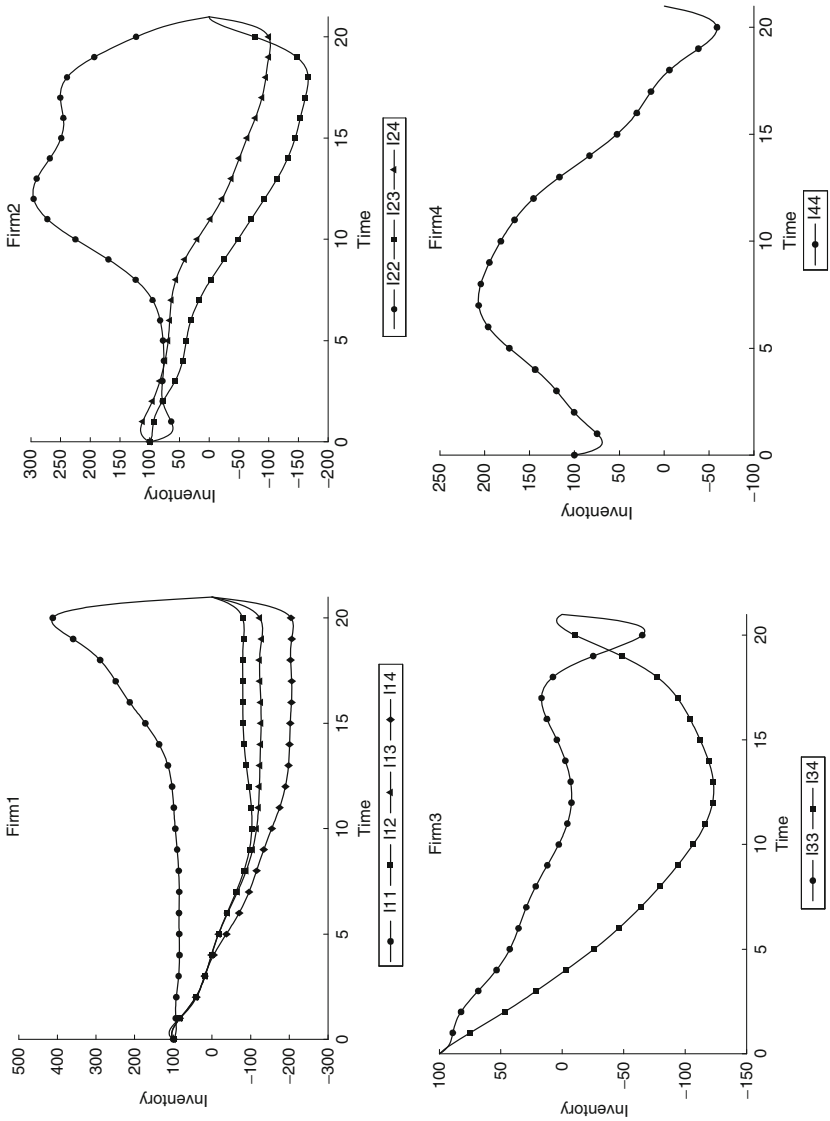


Fig. 8.8 Inventory trajectories (grouped by firms)

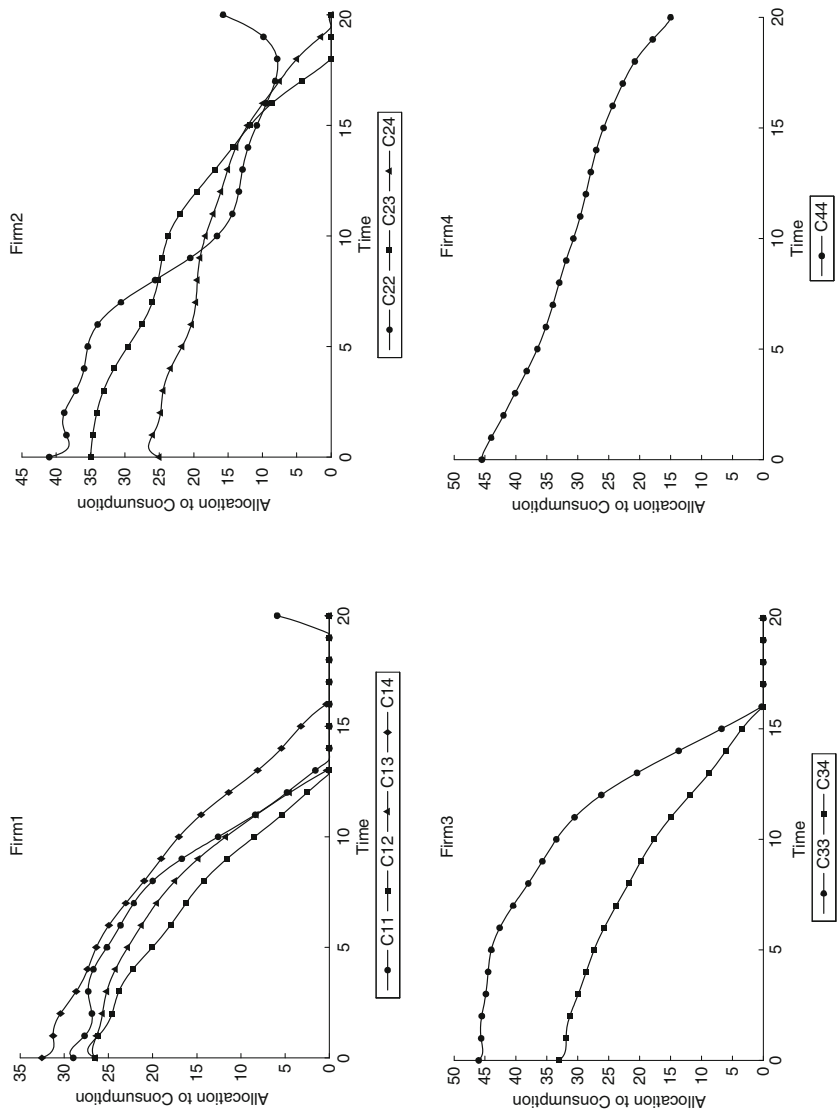
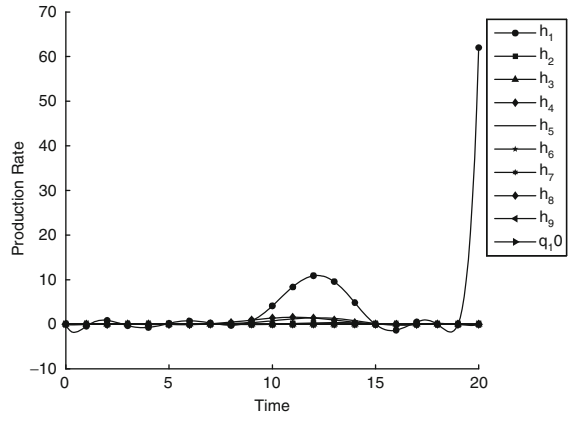
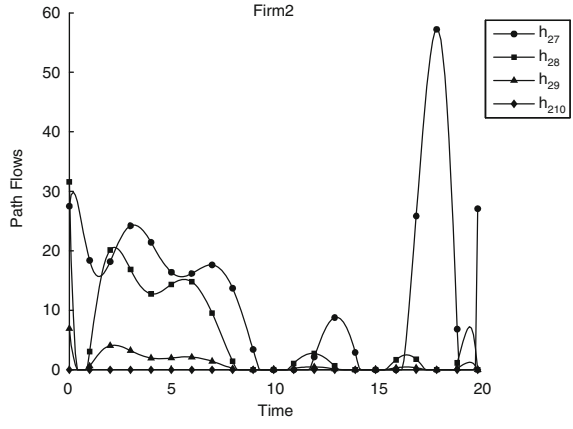


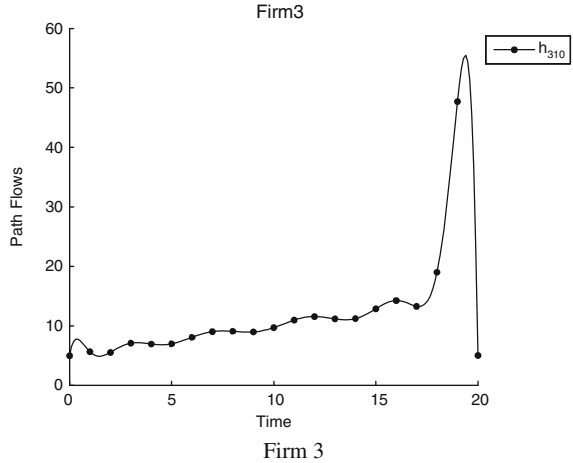
Fig. 8.9 Allocations of output to meet demands



Firm 1



Firm 2



Firm 3

Fig. 8.10 Path flows (grouped by firms)

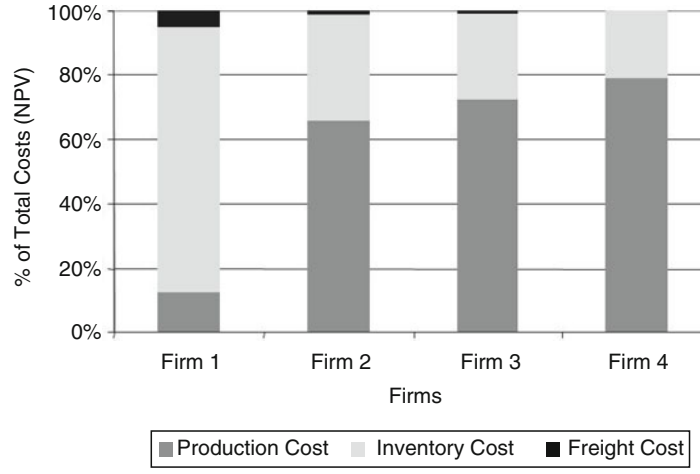


Fig. 8.11 Costs by firm

8.5 Competitive Supply Chains

It is possible to modify and extend the spatial oligopolistic competition model presented in the previous section to include consideration of supply chains. We will use the same notation as that used in Section 8.4, modified to distinguish among producers, suppliers, and retailers. The setting will be one for which a single homogenous good is manufactured by producers for sale by retailers; shipments from producers to retailers are free onboard; there is a multi-echelon supply chain that prepares the input flows to producers and may create and exploit inventories at each level. Additionally, producers may also build up and spatially redistribute inventories to their advantage. In the discussion that follows, it will simplify our notation to imagine all producers and all retailers maintain a presence at every node of the underlying graph. In recognition of this feature, we allow transport between all pairs (i, j) of nodes, where $i, j \in \mathcal{N}$ and $\mathcal{N}$ is the complete set of network nodes. This generality is merely a notational convenience; a model in which individual producers and retailers are restricted to subsets of nodes may be easily created.

8.5.1 Inverse Demands

Agreements are in place that prevent producers from selling directly to consumers, so retailers and consumers will face subtlely distinct inverse demand functions for the finished good of interest. To understand this, let $\mathcal{F}_P$ be the set of producers and $\mathcal{F}_R$ the set of retailers, potentially occupying every network node. Also let w_j refer to the wholesale price paid by retailers $r \in \mathcal{F}_R$ for the producers' ouput in market

$j \in \mathcal{N}$. We note the following identity holds:

$$D_j(w_j + \alpha_j w_j) = \sum_{r \in \mathcal{F}_R} c_j^r \tag{8.69}$$

where $D_j(.)$ is the market demand function at node $j \in \mathcal{N}$ for the finished good and c_j^r is the consumption at node $j \in \mathcal{N}$ of the goods flow from retailer $r \in \mathcal{F}_R$; furthermore, $\alpha_j \in \Re^1_{++}$ is the retailers' margin at node $j \in \mathcal{N}$. We assume that each such demand function has an inverse denoted by $\Theta_j(.,.)$ such that

$$w_j = \Theta_j \left(\sum_{r \in \mathcal{F}_R} c_j^r, \alpha_j \right) \quad \forall j \in \mathcal{N} \tag{8.70}$$

which we call the wholesalers' inverse demand. Next we denote the consumers' inverse demand for the finished good at node $j \in \mathcal{N}$ by $\Psi_j(.)$; that inverse is obtained from (8.69) but is not identical to the wholesalers' inverse demand (8.70). In particular, the consumers' inverse demand takes the form

$$w_j + \alpha_j w_j = v_j = \Psi_j \left(\sum_{r \in \mathcal{F}_R} c_j^r \right) \quad \forall j \in \mathcal{N} \tag{8.71}$$

and $w_j + \alpha_j w_j$ is the retail price paid by consumers for the finished good at node $j \in \mathcal{N}$. We assume that the inverse demand $\Psi_j(.)$ exists for every retail market $j \in \mathcal{N}$. Clearly an alternative form of (8.71) is

$$w_j = \frac{1}{1 + \alpha_j} \Psi_j \left(\sum_{r \in \mathcal{F}_R} c_j^r \right) \quad \forall r \in \mathcal{F}_R, \; j \in \mathcal{N} \tag{8.72}$$

Expressions (8.70) and (8.72) make very clear that

$$\Theta_j \left(\sum_{r \in \mathcal{F}_R} c_j^r, \alpha_j \right) = \frac{1}{1 + \alpha_j} \Psi_j \left(\sum_{r \in \mathcal{F}_R} c_j^r \right) \quad \forall j \in \mathcal{N} \tag{8.73}$$

8.5.2 *Producers' Extremal Problem*

To facilitate the story begun above, we employ the following state dynamics for producers:

$$\frac{dI_i^f}{dt} = F_i^f(h_i^f) + \sum_{j \in \mathcal{N}} s_{ji}^f - \sum_{j \in \mathcal{N}} s_{ij}^f - \sum_{r \in \mathcal{F}_R} \sum_{j \in \mathcal{N}} q_{ij}^{fr} \quad \forall f \in \mathcal{F}_P, \forall r \in \mathcal{F}_R, i \in \mathcal{N} \tag{8.74}$$

where $F_i^f(.)$ is a node-specific single factor production function for producer $f \in \mathcal{F}_P$, and $\mathcal{F}_P$ is the set of firms producing the homogeneous product of interest. In addition, h_i^f is the flow of the single factor to firm $f \in \mathcal{F}_P$ at node $i \in \mathcal{N}$. As already mentioned, the set of retailers we consider is $\mathcal{F}_R$. Additionally, q_{ij}^{fr} denotes the sales by producer $f \in \mathcal{F}_P$ from inventory or new production at node $i \in \mathcal{N}$ to retailer $r \in \mathcal{F}_R$ at node $j \in \mathcal{N}$. Although we refer to the input to production as a single factor, it is in fact a precisely constitutued aggregate of several inputs, constructed from individual factors added at each level of the supply chain through which it passes. Furthermore, N is the final echelon (stage) of supplying the single factor to producers at nodes $i \in \mathcal{N}$. Moreover, the aggregate input factor flow u_N is disaggregated into individual flows h_i^f used by each firm $f \in \mathcal{F}_P$ at each node $i \in \mathcal{N}$ where

$$u_N = \sum_{f \in \mathcal{F}_P} \sum_{i \in \mathcal{N}} h_i^f \tag{8.75}$$

Naturally we employ the notation

$$h^f = (h_i^f : i \in \mathcal{N}) \tag{8.76}$$

to describe the vector of factor allocations controlled by producer $f \in \mathcal{F}_P$. We assume there are supply contracts in place, between producer $f \in \mathcal{F}_P$ and the supply chain agent, that have the effect of establishing a cost function $C_i^f(h_i^f, t)$ for the instantaneous cost to acquire the input flow h_i^f at each node $i \in \mathcal{N}$. Additionally, the factor flow to each producer is constrained according to

$$A_f \leq \sum_{i \in \mathcal{N}} h_i^f \leq B_f \quad \forall f \in \mathcal{F}_P \tag{8.77}$$

In (8.77)

$$A_f, B_f \in \Re_{++}^1$$

are the lower and upper bounds (on factor flows) established by the aformentioned contracts.

In light of the above development, we may express the extremal problem for each producer $f \in \mathcal{F}_P$ as follows:

$$\begin{aligned} \max J_P^f(q^f, s^f, h^f; c) = \int_{t_0}^{t_f} e^{-\rho t} & \sum_{r \in \mathcal{F}_R} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \frac{1}{1+\alpha_j^r} \Psi_j \left(\sum_{g \in \mathcal{F}_P} c_j^g \right) q_{ij}^{fr} \\ & - \sum_{i \in \mathcal{N}} C_i^f \left(h_{iN}^f, t \right) - \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} r_{ij} \left(s_{ij}^f + \sum_{r \in \mathcal{F}_R} q_{ij}^{fr} \right) \\ & - \sum_{i \in \mathcal{N}} \psi_i^f \left(I_i^f, t \right) \Bigg\} dt \end{aligned} \tag{8.78}$$

subject to

$$\frac{dI_i^f}{dt} = F_i^f(h_i^f) + \sum_{j\in\mathcal{N}} s_{ji}^f - \sum_{j\in\mathcal{N}} s_{ij}^f - \sum_{r\in\mathcal{F}_R}\sum_{j\in\mathcal{N}} q_{ij}^{fr} \quad \forall f \in \mathcal{F}_P, \forall r \in \mathcal{F}_R, i \in \mathcal{N} \tag{8.79}$$

and

$$I_i^f(t_0) = K_i^f \quad \forall i \in \mathcal{N} \tag{8.80}$$

$$I_i^f(t_f) \geq \tilde{K}_i^f \quad \forall i \in \mathcal{N} \tag{8.81}$$

$$0 \leq q_{ij}^{fr} \leq \widetilde{q}_{ij}^{fr} \quad i, j \in \mathcal{N}, r \in \mathcal{F}_r \tag{8.82}$$

$$0 \leq s_{ij}^f \leq \widetilde{s}_{ij}^f \quad i, j \in \mathcal{N} \tag{8.83}$$

$$A_f \leq \sum_{i\in\mathcal{N}} h_i^f \leq B_f \quad \forall f \in \mathcal{F}_P \tag{8.84}$$

$$u_N = \sum_{g\in\mathcal{F}_P}\sum_{i\in\mathcal{N}} h_{Ni}^g \tag{8.85}$$

where q^f, s^f, and h^f are vectors, which are vectors of output, shipping, and input factor flows under the control of producer $f \in \mathcal{F}_P$. We recall that shipments to retailers are free onboard and thus paid by producers. Furthermore, r_{ij} is the freight rate for origin-destination (OD) pair (i, j), while inventory cost for producer $f \in \mathcal{F}_P$ at node $i \in \mathcal{N}$ is $\psi_i^f(.,.)$; also K_i^f is the initial inventory for producer $f \in \mathcal{F}_P$ at node $i \in \mathcal{N}$. Moreover, (8.81) is the terminal time inventory constraint; (8.82) and (8.83) are constraints expressing bounds on outputs and shipments, where $\tilde{K}_i^f$, $\widetilde{q}_{ij}^{fr}$, and $\widetilde{s}_{ij}^f$ are exogenous fixed parameters for all $f \in \mathcal{F}_P$ and $i, j \in \mathcal{N}$. Constraints (8.84) form the aforementioned upper and lower bounds on aggregate factor flows to producers. Note that

$$q_{ij}^f = \left(q_{ij}^{fr} : r \in \mathcal{F}_R\right) \tag{8.86}$$

$$q^f = (q_{ij}^f : i, j \in \mathcal{N}) \tag{8.87}$$

$$q^{-f} = (q^g : g \in \mathcal{F}_P \backslash f) \tag{8.88}$$

$$s^f = \left(s_{ij}^f : i, j \in \mathcal{N}\right) \tag{8.89}$$

$$h^f = (h_i^f : i \in \mathcal{N}) \tag{8.90}$$

$$h = \left(h^f : f \in \mathcal{F}_P\right) \tag{8.91}$$

$$h^{-f} = (h^g : g \in \mathcal{F}_P \backslash f) \tag{8.92}$$

$$c^r = \left(c_j^r : j \in \mathcal{N}\right) \tag{8.93}$$

$$c = (c^r : r \in \mathcal{F}_R) \tag{8.94}$$

Furthermore, the constraints of this extremal problem depend on the vector h^{-f} and on the scalar unknown u_N, both of which are determined exogenously by the producers and the supply chain manager, respectively.

8.5.3 Retailers' Extremal Problem

Turning our attention to retailers, we stipulate that only retailers may sell finished goods. Since the single homogeneous finished good must be obtained from producers, the pertinent dynamics for retailers are

$$\frac{dR_j^r}{dt} = \sum_{f \in \mathcal{F}_P} \sum_{i \in \mathcal{N}} q_{ij}^{fr} - c_j^r \qquad \forall r \in \mathcal{F}_R,\ j \in \mathcal{N} \tag{8.95}$$

where R_j^r denotes the inventory of retailer $r \in \mathcal{F}_R$ at node $j \in \mathcal{N}$, while $\mathcal{F}_R$ is the set of retailers and $\mathcal{N}$ is the set of nodes at which retailer r is located. Note also that c_j^r denotes the consumption activity served by retailer $r \in \mathcal{F}_R$ at node $j \in \mathcal{N}$. Therefore, the extremal problem faced by each retailer $r \in \mathcal{F}_R$ is the following:

$$\begin{aligned} \max J_R^r(c^r, q^r; c^{-r}) = & \int_{t_0}^{t_f} e^{-\rho t} \sum_{j \in \mathcal{N}} \left(c_j^r - \frac{1}{1+\alpha_j^r} \sum_{i \in \mathcal{N}} q_{ij}^{fr} \right) \Psi_j \left(\sum_{g \in \mathcal{F}_P} c_j^g \right) \\ & - \sum_{j \in \mathcal{N}} \phi_j^r(R_j^r, t) dt \end{aligned} \tag{8.96}$$

subject to

$$\frac{dR_j^r}{dt} = \sum_{f \in \mathcal{F}_P} \sum_{i \in \mathcal{N}} q_{ij}^{fr} - c_j^r \qquad \forall r \in \mathcal{F}_R,\ j \in \mathcal{N} \tag{8.97}$$

$$0 \le c_j^r \le \widetilde{c}_j^r \qquad \forall\, r \in \mathcal{F}_R, j \in \mathcal{N} \tag{8.98}$$

$$R_j^r(t_0) = Q_j^r \qquad \forall\, r \in \mathcal{F}_R, j \in \mathcal{N} \tag{8.99}$$

$$R_j^r(t_f) = \widetilde{Q}_j^r \qquad \forall\, r \in \mathcal{F}_R, j \in \mathcal{N} \tag{8.100}$$

In (8.96), $\phi_i^r\left(R_j^r, t\right)$ denotes the inventory costs at node $i \in \mathcal{N}$ for retailer $r \in \mathcal{F}_R$. Additionally, Q_j^r is the initial inventory and $\widetilde{Q}_j^r$ is the terminal time inventory, while $\widetilde{c}_j^r$ is the upper bound on consumption, for retailer $r \in \mathcal{F}_R$ at node $i \in \mathcal{N}$. Note that

$$c^r = \left(c_j^r : j \in \mathcal{N}\right) \tag{8.101}$$

$$c^{-r} = (c^g : g \in \mathcal{F}_P - \{r\}) \tag{8.102}$$

$$q_{ij}^r = \left(q_{ij}^{fr} : f \in \mathcal{F}_P\right) \tag{8.103}$$

$$q^r = \left(q_{ij}^r : i, j \in \mathcal{N}\right) \tag{8.104}$$

Note that the constraints of this extremal problem depend on the vector q^r, which is exogenous, since output allocations are determined by the producers.

8.5.4 Supply Chain Extremal Problem

Now let us consider a multi-echelon supply chain stretching from unrefined raw materials to factor flows ready for use by producers. We use u_k to denote the flow of the input factor exiting stage k (i.e., the flow from stage k to stage $k+1$). If we use S_k to denote the inventory at stage k of the supply chain, we may write

$$\frac{dS_k}{dt} = u_{k-1} - u_k \quad k = 1, \ldots, N$$

where it is understood that only the terminal flow u_N is ready for use in producing the homogeneous finished good of present interest to us. Recall that we have already assumed there are contracts in place specifying a fee schedule $C_i^f(h_i^f, t)$ and guarantied upper and lower bounds for factor flow to each producer $f \in \mathcal{F}_p$ at node $i \in \mathcal{N}$ at time $t \in \left[t_0, t_f\right]$, where the allocations h^f are controlled by producer $f \in \mathcal{F}_P$. The controls available to the supply chain agent are captured by the vector

$$u = (u_k : k \in [1, N]) \tag{8.105}$$

As a consequence the single manager who operates all supply chain stages $k \in [1, N]$ seeks to minimize his/her total cost; that is, he/she seeks to solve the following optimal control problem:

$$\min J_S(u) = \int_{t_0}^{t_f} e^{-\rho t} \sum_{k=1}^{N} \left[V_k(u_k, t) + \varphi_k(S_k, t)\right] dt \tag{8.106}$$

subject to

$$\frac{dS_k}{dt} = u_{k-1} - u_k \quad k \in [1, N] \tag{8.107}$$

$$S_k(0) = S_k^0 \qquad k \in [1, N] \tag{8.108}$$

$$u_N = \sum_{f \in \mathcal{F}_P} \sum_{i \in \mathcal{N}} h_i^f \tag{8.109}$$

$$0 \leq u_k \leq \tilde{u}_k \quad k \in [1, N] \tag{8.110}$$

where $V_k(.,.)$ denotes the variable costs of preparing the stage k flow, $\varphi_k(.,.)$ is the inventory cost function, S_k^0 is the initial inventory, and $\tilde{u}_k$ is the technological upper bound for stage k flow of the supply chain. Note that constraint (8.109) was introduced prevously as (8.75). Note also that the constraints of this extremal problem depend on the vector h which is determined exogenously, since the producers decide factor flows to their production facilities consistent with the contracts they hold with the supply chain manager.

8.5.5 The Differential Variational Inequality

In this section we give an overview of how the relevant differential variational inequality for our combined producer-retailer-supply chain game may be formed.

8.5.5.1 Maximum Principle for the Producers

With

$$\begin{pmatrix} c \\ u_N \\ q^{-f} \end{pmatrix} \tag{8.111}$$

as exogenous, each producer $f \in F_P$ solves

$$\max J_P^f(q^f, s^f, h^f; c) \quad \text{s.t.} \quad \left(q^f, s^f, h^f\right) \in \Lambda_P^f(h^{-f}, u_N) \tag{8.112}$$

where

$$\Lambda_P^f(h^{-f}, u_N) \equiv \left\{ \begin{pmatrix} q^f \\ s^f \\ h^f \end{pmatrix} : (8.79), (8.80), (8.81), (8.82), \right.$$
$$\left. (8.83), (8.84), (8.85), \text{ adjoint equations, and transversality hold} \right\} \tag{8.113}$$

The corresponding Hamiltonian is

$$H_P^f(q^f, s^f, h^f, I^f, \lambda^f; c) = e^{-\rho t} \left\{ \sum_{r \in \mathcal{F}_R} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \frac{1}{1+\alpha_j^r} \Psi_j \left(\sum_{g \in \mathcal{F}_P} c_j^g \right) q_{ij}^{fr} \right.$$
$$\left. - \sum_{i \in \mathcal{N}} C_i^f(h_{iN}^f, t) - \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} r_{ij} \left(s_{ij}^f + \sum_{r \in \mathcal{F}_R} q_{ij}^{fr} \right) \right.$$

$$- \sum_{i \in \mathcal{N}} \psi_i^f (I_i^f, t) \Bigg\}$$
$$+ \sum_{i \in \mathcal{N}} \lambda_i^f \left\{ F_i^f (h_i^f) + \sum_{j \in \mathcal{N}} s_{ji}^f - \sum_{j \in \mathcal{N}} s_{ij}^f - \sum_{r \in \mathcal{F}_R} \sum_{j \in \mathcal{N}} q_{ij}^{fr} \right\}$$

where $I^f = (I_i^f : i \in \mathcal{N})$ and $\lambda^f = (\lambda_i^f : i \in \mathcal{N})$ is a vector of adjoint variables. We will use the notation

$$\nabla_{z^f} H_P^{f*} = \nabla_z H_P^f (q^{f*}, s^{f*}, h^{f*}, I^{f*}, \lambda^{f*}; c^*) \tag{8.114}$$

to denote the gradient of the Hamiltonian with respect to the control vector

$$z^f = \begin{pmatrix} q^f \\ s^f \\ h^f \end{pmatrix}$$

of producer f evaluated at a Nash equilibrium. The maximum principle for producer $f \in \mathcal{F}_P$ leads to:

$$\left[\nabla_{q^f} H_P^{f*}\right]^T (q^f - q^{f*}) + \left[\nabla_{s^f} H_P^{f*}\right]^T (s^f - s^{f*}) + \left[\nabla_{h^f} H_P^{f*}\right]^T (h^f - h^{f*}) \leq 0 \tag{8.115}$$

$$\begin{pmatrix} q^f \\ s^f \\ h^f \end{pmatrix}, \begin{pmatrix} q^{f*} \\ s^{f*} \\ h^{f*} \end{pmatrix} \in \Lambda_P^f (h^{-f*}, u_N^*) \tag{8.116}$$

8.5.5.2 *Maximum Principle for the Retailers*

With

$$\begin{pmatrix} c^{-r} \\ q^r \end{pmatrix} \tag{8.117}$$

as exogenous, each retailer $r \in \mathcal{F}_R$ solves

$$\max J_R^r (c^r, q^r; c^{-r}) \quad \text{s.t.} \quad c^r \in \Lambda_R^r (q^r) \tag{8.118}$$

where

$$\Lambda_R^r (q^r) \equiv \{c^r : (8.97), (8.98), (8.99), (8.100), \text{adjoint equations, and transversality hold}\} \tag{8.119}$$

The corresponding Hamiltonian is

$$H_R^r(c^r, R^r, \gamma^r; c^{-r}, q^r) = e^{-\rho t} \left[\sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}} \left(c_j^r - \frac{1}{1+\alpha_j^r} \sum_{i\in\mathcal{N}} q_{ij}^{fr} \right) \Psi_j \left(\sum_{g\in\mathcal{F}_R} c_j^g \right) - \sum_{j\in\mathcal{N}} \phi_j^r(R_j^r, t) \right] + \sum_{j\in\mathcal{N}} \gamma_j^r \left(\sum_{f\in\mathcal{F}_P} \sum_{i\in\mathcal{N}} q_{ij}^{fr} - c_j^r \right)$$

where $R^r = \left(R_j^r : j \in \mathcal{N}\right)$ and $\gamma^r = \left(\gamma_j^r : j \in \mathcal{N}\right)$ is a vector of adjoint variables. We will use the notation

$$\nabla_{c^r} H_R^{r*} = \nabla_{c^r} H_R^r \left(c^{r*}, R^{r*}, \gamma^{r*}; c^{-r*}, q^{r*}\right) \tag{8.120}$$

to denote the gradient of the Hamiltonian with respect to the controls of retailer r evaluated at a Nash equilibrium. The maximum principle for retailer $r \in \mathcal{F}_R$ leads to:

$$\left[\nabla_{c^r} H_R^{r*}\right]^T \left(c^r - c^{r*}\right) \le 0 \quad c^r, c^{r*} \in \Lambda_R^r \left(q^{r*}\right) \tag{8.121}$$

8.5.5.3 Minimum Principle for the Supply Chain

With h as exogenous, the supply chain manager solves

$$\min J_S(u) \quad \text{s.t.} \quad u \in \Lambda_S(h) \tag{8.122}$$

where

$$h = \left(h^f : f \in \mathcal{F}_P\right) \tag{8.123}$$

and

$$\begin{aligned} \Lambda_S(h) \equiv \{&u : (8.107), (8.108), (8.109), (8.110), \text{adjoint equations,} \\ &\text{and transversality hold}\} \end{aligned} \tag{8.124}$$

The corresponding Hamiltonian is

$$H_S(u, S, \zeta; h) = e^{-\rho t} \sum_{k=1}^{N} \left[V_k(u_k, t) + \varphi_k\left(S_k, t\right)\right] + \sum_{k=1}^{N} \zeta_k \left(u_{k-1} - u_k\right)$$

where $S = (S_k : k \in [i, N])$ and $\zeta = (\zeta_k : k \in [1, N])$ is a vector of adjoint variables. We will use the notation

$$\nabla_u H_S^* = \nabla_u H_S \left(u^*, S^*, \zeta^*; h^*\right) \tag{8.125}$$

to denote the gradient of the Hamiltonian with respect to supply chain controls evaluated at a Nash equilibrium. The minimum principle for the single supply chain manager leads to:

$$\left[\nabla_u H_S^*\right]^T (u - u^*) \geq 0 \quad u, u^* \in \Lambda_S(h^*) \tag{8.126}$$

8.5.6 The DVI

We note that the finite-dimensional variational inequalities derived above hold for each instant of continuous time. So we may integrate the individual variational inequalities over time and sum them over discrete agent indices to obtain a single necessary condition: the solution

$$\begin{pmatrix} q^* \\ s^* \\ h^* \\ c^* \\ u^* \end{pmatrix} \in \Lambda \tag{8.127}$$

must satisfy

$$\begin{aligned} \sum_{f \in \mathcal{F}_P} \int_{t_0}^{t_f} \Big\{ \left[-\nabla_{q^f} H_P^{f*}\right]^T (q^f - q^{f*}) \; + \; \left[-\nabla_{s^f} H_P^{f*}\right]^T (s^f - s^{f*}) \\ + \left[-\nabla_{h^f} H_P^{f*}\right]^T (h^f - h^{f*}) \Big\} \, dt \\ + \int_{t_0}^{t_f} \sum_{r \mathcal{F}_R} \left[-\nabla_{c^r} H_R^{r*}\right]^T (c^r - c^{r*}) dt \\ + \int_{t_0}^{t_f} \left[\nabla_u H_S^*\right]^T (u - u^*) dt \geq 0 \end{aligned} \tag{8.128}$$

for all

$$\begin{pmatrix} q \\ s \\ h \\ c \\ u \end{pmatrix} \in \Lambda(h, q, u) \tag{8.129}$$

where

$$\Lambda = \Lambda_S(h^*) \times \prod_{f \in \mathcal{F}_P} \Lambda_P^f(h^{-f}, u_N) \times \prod_{r \in \mathcal{F}_R} \Lambda_R^r\left(q^{r*}\right) \tag{8.130}$$

8.5.7 Numerical Example

Let us consider the network of Figure 8.12, which has two arcs and three nodes for suppliers and nine arcs and seven nodes for producers and retailers. In addition, there are five arcs between suppliers and producers. Producer 1 and producer 2 have activities located at nodes $i = 1, 2, 3, 4, 5$. Retailer 1 is located at node 6; retailer 2 is located at node 7. The time interval of interest is $[0, 10]$; that is $t_0 = 0$ and $t_f = 10$. The initial inventories at each node for producers and retailers are the following:

$$\begin{array}{lll} I_1(0) = 2 & S_0(0) = 10 & R_6(0) = 1 \\ I_2(0) = 3 & S_1(0) = 3 & R_7(0) = 1 \\ I_3(0) = 2 & S_2(0) = 1 & \\ I_4(0) = 3 & & \\ I_5(0) = 2 & & \end{array}$$

The discount rate ρ is assumed to be 0.05 and the retailers' margins at nodes $j = 6$ and $j = 7$ are 0.1 and 0.08, respectively. Keeping in mind that the retailers occupy

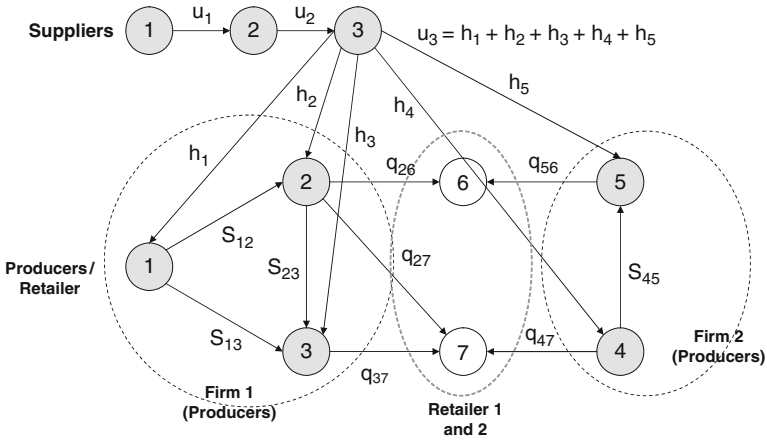


Fig. 8.12 Integrated supply, production and distribution network

distinct nodes in this example, we may assume the consumers' inverse demand functions for nodes $i = 6$ and $i = 7$ are the following:

$$\Psi_6(c_6, c_7) = 11 - c_6$$
$$\Psi_7(c_6, c_7,) = 11 - 1.5 \cdot c_7$$

where c_j is the consumption at j from the sole retailer located there. The production cost functions at nodes $i = 1, 2, 3, 4, 5$ are the following:

$$\begin{array}{ll} F_1(h_1) = 0.50\,(h_1)^2 & F_4(h_4) = 0.20\,(h_4)^2 \\ F_2(h_2) = 0.10\,(h_2)^2 & F_5(h_5) = 0.30\,(h_5)^2 \\ F_3\,(h_3) = 0.15\,(h_3)^2 & \end{array}$$

where we use h_i to denote the input factor flow from the final (third) stage of the supply chain to producer i. The inventory cost functions for producers $i = 1, 2, 3, 4, 5$ are

$$\begin{array}{ll} \psi_1(I_1, t) = 1.5\,(h_1)^2 & \psi_4(I_4, t) = 2.0\,(h_4)^2 \\ \psi_2(I_2, t) = 1.5\,(h_2)^2 & \psi_5(I_5, t) = 1.5\,(h_5)^2 \\ \psi_3(I_3, t) = 1.5\,(h_3)^2 & \end{array}$$

The inventory cost functions for retailers at nodes $i = 6, 7$ are

$$\phi_6(R_6, t) = 3.0\,(R_6)^2 \qquad \phi_7(R_7, t) = 2.5\,(R_7)^2$$

The inventory cost functions for the $k = 1, 2, 3$ stages of the supply chain are

$$\varphi_1(S_1, t) = 5.0\,(S_1)^2 \;\; \varphi_2(S_2, t) = \frac{3}{8}\,(S_2)^2 \;\; \varphi_3(S_3, t) = 1.0\,(S_3)^2$$

We assume that unit freight costs between nodal pairs are constant:

$$\begin{array}{lll} r_1 = 2.0 & r_4 = 0.5 & r_7 = 5.0 \\ r_2 = 2.0 & r_5 = 4.0 & r_8 = 3.0 \\ r_3 = 0.5 & r_6 = 3.0 & r_9 = 4.0 \end{array}$$

The costs to producers at $i = 1, 2, 3, 4, 5$ for acquiring factor input flows at $i = 1, 2, 3, 4, 5$ are the following:

$$\begin{array}{ll} C_1(h_1, t) = 2.0\,(h_1) & C_4(h_4, t) = 3.0\,(h_4) \\ C_2(h_2, t) = 2.5\,(h_2) & C_5(h_5, t) = 2.1\,(h_5) \\ C_3(h_3, t) = 2.1\,(h_3) & \end{array}$$

Finally, the variable costs of preparing the stage $k = 1, 2, 3$ supply chain flows are

$$V_1(u_1, t) = 1.2\,(u_1) \qquad V_2(u_2, t) = 1.3\,(u_2) \qquad V_3(u_3, t) = 2.0\,(u_3)$$

The upper bounds on control variables are

$$\widetilde{q} = \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} \quad \tilde{s} = \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} \; B_f = \begin{pmatrix} 8 \\ 8 \\ 8 \\ 8 \\ 8 \end{pmatrix} \; \widetilde{u} = \begin{pmatrix} 15 \\ 15 \\ 15 \end{pmatrix} \; \widetilde{c} = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$$

Keeping in mind that the subnetworks for producers and retailers are disjoint, the initial-value problems that consitute inventory dynamics for producers $i = 1, 2, 3, 4, 5$ are the following flow balance equations:

$$\frac{dI_1}{dt} = F_1(h_1) - s_{12} - s_{13} \qquad I_1(t_0) = I_1^0 = 2$$

$$\frac{dI_2}{dt} = F_2(h_2) + s_{12} - s_{23} - q_{26} \qquad I_2(t_0) = I_2^0 = 3$$

$$\frac{dI_3}{dt} = F_3(h_3) + s_{13} + s_{23} - q_{37} \qquad I_3(t_0) = I_3^0 = 2$$

$$\frac{dI_4}{dt} = F_4(h_4) - s_{45} - q_{46} \qquad I_4(t_0) = I_4^0 = 3$$

$$\frac{dI_5}{dt} = F_5(h_5) + s_{45} - q_{56} \qquad I_5(t_0) = I_5^0 = 2$$

where q_{ij} is the flow from a producer at i to a retailer at j.

Inventory dynamics for retailers $i = 6, 7$ are the following flow balance equations

$$\frac{dR_6}{dt} = q_{26} + q_{56} - c_6 \qquad R_6(t_0) = R_6^0 = 1$$

$$\frac{dR_7}{dt} = q_{27} + q_{37} + q_{46} - c_7 \qquad R_7(t_0) = R_7^0 = 1$$

Inventory dynamics for the $k = 1, 2, 3$ supply chain stages are the following flow balance equations:

$$\frac{dS_1}{dt} = u_0 - u_1 \qquad S_1(t_0) = S_1^0 = 10$$

$$\frac{dS_2}{dt} = u_1 - u_2 \qquad S_2(t_0) = S_2^0 = 3$$

$$\frac{dS_3}{dt} = u_2 - u_3 \qquad S_2(t_0) = S_2^0 = 1$$

Furthermore, we have

$$u_3 = h_1 + h_2 + h_3 + h_4 + h_5$$

We leave, as an exercise for reader, the formulation of the relevant Hamiltonians and the encompassing differential variational inequality. In this numerical example, we employ production functions exhibiting increasing returns to scale, as well as linear demand functions. A fixed-point algorithm implemented in discrete time is an attractive numerical scheme since the dynamics are linear. Note, however, that the model considered is actually a differential quasivariational inquality; thus, its discrete time counterpart, is a finite-dimensional quasivariational inequality. As such the unembellished fixed-point algorithm may fail to converge, since a proof of its convergence for general quasi-variational inequalities is not available. That is, the fixed-point algorithm is a heuristic in the present application. A complete presentation of this example's mathematical formulation, associated data, algorithmic details and numerical solution may be found by following self-explanatory links found at the website http://www2.ie.psu.edu/csee/DODG/Ch8NuEx.pdf.

8.6 Exercises

1. Study some of the consumption, inventory, and shipping trajectories from the models of this chapter, paying particular attention to what happens as the terminal time is approached. (a) Can you explain any of the phenomena observed? (b) In particular, are implausible actions attributed to a producer by a production planning model when a finite horizon is employed without requiring that inventory vanish when the horizon is reached? (c) How can one compute steady states of the models of this chapter given the complexity of observed behavior near the terminal time?
2. Using necessary and sufficient conditions, articulate decision rules for the models presented in this chapter. How would these decision rules be employed in practice?
3. For each of the models expressed as a differential variational inequality (DVI) in this chapter, provide conditions that assure any solution of the DVI is also a solution of the differential Nash game being studied.
4. For each of the models expressed as a differential variational inequality (DVI) in this chapter, provide conditions that ensure a solution exists.
5. Develop a supply chain model like that of Section 8.5, but with a single, monopolistic producer of the finished goods delivered to retailers.
6. State regularity conditions that will assure the existence and uniqueness of state operators for the supply chain model of Section 8.5.
7. State relevant regularity conditions and provide a proof of the existence of a solution to the supply chain model of Section 8.5.

8. Give the complete statement of the feasible region of the supply chain differential variational inequality appearing in expression (8.130).
9. Create a small two-firm, two-market numerical example of the spatial oligopolistic network competition model presented in Section 8.4. Solve it by using the maximum principle.
10. For the dynamic aspatial monopoly considered in Section 8.1, reformulate the problem for noninvertible, nonseparable demand functions. Analyze and give an economic interpretation of the optimality conditions. What difficulties do you foresee in the numerical solution of your reformulation? Have we developed any algorithms in previous chapters that apply?

List of References Cited and Additional Reading

Anderson, W. H. L. (1970). Production scheduling, intermediate goods, and labor productivity. *American Economic Review 60*, 153–162.

Arrow, K. J. and S. Karlin (1958). Production over time with increasing marginal costs. In K. J. Arrow, S. Karlin, and H. Scarf (Eds.), *Studies in the Mathematical Theory of Inventory and Production*. Palo Alto: Stanford University Press.

Friesz, T. L., M. A. Rigdon, and R. Mookherjee (2006). Differential variational inequalities and shipper dynamic oligopolistic network competition. *Transportation Research Part B 40*, 480–503.

Graves, S., D. Kletter, and W. Hetzel (1998). A dynamic model for requirements planning with application to supply chain optimization. *Operations Research 46*(3), S35–S49.

Lieber, Z. (1973). An extension to Modigliani and Hohn's planning horizons results. *Management Science 20*, 319–330.

Morin, F. (1955). Note on an inventory problem discussed by Modigliani and Hohn. *Econometrica 23*, 447–452.

Nagurney, A., J. Dong, and D. Zhang (2002). A supply chain network equilibrium model. *Transportation Research Part E 38*(5), 281–303.

Pekelman, D. (1974). Simultaneous price production decisions. *Operations Research 22*, 788–794.

Sethi, S. P. and G. L. Thompson (1981). *Optimal Control Theory: Applications to Management Science*. Boston: Martinus Nijhoff.

Smith, V. L. (1961). *Invenstment and Production: A Study in the Theory of the Capital Using Enterprise*. Cambridge, MA: Harvard University Press.

Sprzeuzkouski, A. Y. (1967). A problem in optimal stock management. *Journal of Optimization Theory and Applications 1*, 232–241.

Swaminathan, J., S. Smith, and N. Sadeh (1998). Modeling supply chain dynamics: a multiagent approach. *Decision Sciences 29*(3), 607–632.

Talluri, S. and R. Baker (2002). A multiphase mathematical programming approach for effective supply chain design. *European Journal of Operational Research 141*(3), 544–558.

Chapter 9
Dynamic User Equilibrium

Dynamic traffic assignment (DTA) is the positive (descriptive) modeling of time-varying flows of automobiles on road networks consistent with established traffic flow theory and travel demand theory. Dynamic user equilibrium (DUE) is one type of DTA wherein the effective unit travel delay, including early and late arrival penalties, of travel for the same purpose is identical for all utilized path and departure time pairs. In the context of planning, DUE is usually modelled for the within-day time scale based on demands established on a day-to-day time scale.

In the last several years, much effort has been expended to develop a theoretically sound formulation of dynamic network user equilibrium that is also a canonical form acceptable to scholars and practitioners alike. DUE models tend to be comprised of four essential submodels:

1. a model of path delay;
2. flow dynamics;
3. flow propagation constraints; and
4. a path/departure-time choice model.

Peeta and Ziliaskopoulos (2001), in a comprehensive review of DTA and DUE research, note that there are several published models comprised of the four submodels named above.

We are interested in this chapter in investigating how such DUE models may be mathematically characterized and numerically solved using the notion of a differential variational inequality in Hilbert space to capture their game-theoretic nature. To that end we focus on two infinite-dimensional variational inequality formulations of the DUE problem reported in Friesz et al. (1993) and Friesz et al. (2001) that have much in common with other published models. In fact, the Friesz et al. (1993) and Friesz et al. (2001) formulations are more computationally demanding than most if not all other DUE models because of the complicated path delay operators, equations of motion, and time lags they embody. As such, the algorithms discussed in this chapter should work as well or better when adapted to other DUE models, including those for which path delay is determined by a nonlinear response surface or by simulation in conjunction with a so-called rolling horizon.

In the next section, we review four categories of arc dynamics and associated flow propagation constraints in order to motivate the model formulation presented

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_9, © Springer Science+Business Media, LLC 2010

in this chapter. This review focuses on those antecedent efforts that are the most useful in motivating our approach. After this literature review, we derive flow propagation constraints. This is followed by a mathematical statement of the problem of finding a dynamic user equilibrium (DUE), on a network, as a differential variational inequality. We also provide a formal demonstration that any solution of it is a dynamic user equilibrium relative to both departure time and path choice. A discussion of algorithms and a numerical example conclude this chapter.

The following is a preview of the principal topics covered in this chapter:

Section 9.1: Some Background. In this section, some general remarks are made and some terminology introduced that will helpful in the remaining sections.

Section 9.2: Arc Dynamics. In this section, we review some alternative arc dynamics and select the dynamics we will emphasize in this chapter.

Section 9.3: The Measure-Theoretic Nature of DUE. In this section, we explain why the mathematical expression of a dynamic user equilibrium (DUE) intrinsically requires a measure-theoretic perspective.

Section 9.4: The Infinite-Dimensional Variational Inequality Formulation. In this section, we present an infinite-dimensional variational inequality formulation of the DUE.

Section 9.5: When Delays Are Exogenous. In this section, DUE is investigated for the circumstance of exogenous effective path delays.

Section 9.6: When Delays Are Endogenous. In this section, DUE is investigated for the circumstance of endogenous effective path delays.

9.1 Some Background

The Friesz et al. (1993) formulation is an exact formulation of dynamic network user equilibrium, where by "exact" we mean a model that is completely mathematically internally consistent and involves no *ad hoc* treatment of delay operators, departure time choice, flow propagation anomalies, or other critical model features prior to its numerical solution. The Friesz et al. (1993) formulation employs path delay operators that obey appropriate arc dynamics and incorporate a path flow propagation mechanism. This embedded path flow propagation mechanism ensures arc entry and exit at appropriate times along a given path and preserves the first-in-first-out (FIFO) queue discipline when appropriate regularity conditions are met. The requirement that the delay operators reflect arc level dynamics and flow propagation considerations makes the delay operators unknowable in closed form. The flow propagation mechanism of the Friesz et al. (1993) formulation depends on arc exit time functions and their inverses. Inverse exit time functions, like the path delay operators, cannot be known in closed form. The Friesz et al. (1993) model

expresses a dynamic Nash-like equilibrium relative to departure time and path choice as an infinite-dimensional variational inequality. This variational inequality cannot be solved by traditional methods since it is based on nonanalytic path delay operators which are only known numerically. In subsequent papers, Friesz et al. (2001) and Friesz and Mookherjee (2006) developed a differential variational inequality formulation of dynamic user equilibrium, equivalent to the Friesz et al. (1993) formulation.

9.2 Arc Dynamics

It is possible to write the dynamics of flow on arcs of a network in different ways. In this section, we explore dynamics that view the rate of change of arc volume as equal to the difference between entrance flow and exit flow at each instant of time. To form arc dynamics of this type, several perspectives, which we next review, have been proposed in the literature.

9.2.1 Dynamics Based on Arc Exit Flow Functions

If one posits that it is possible to specify and empirically estimate, or to mathematically derive from some plausible theory, functions that describe the rate at which traffic exits a given network arc for any given volume of traffic present on that arc, one is led to some deceptively simple traffic dynamics. To express this supposition symbolically, we use $x_a(t)$ to denote the volume of traffic on arc a at time t and $g_a(x_a(t))$ to denote the rate at which traffic exits from link a. Where it will not be confusing, we suppress the explicit reference to time t and write the arc volume as x_a and the exit flow function as $g_a(x_a)$ with the understanding that both entities are time varying. It is also necessary to define the rate at which traffic enters arc a, which we denote as $u_a(t)$. Again, when it is not confusing, we may suppress the time dependency of the entrance rate for arc a and simply write u_a. Note that both $g_a(x_a)$ and u_a are rates; that is, they have the units of volume per unit time, so it is appropriate to refer to them as exit flow and entrance flow, respectively. A natural flow balance equation can now be written for each link:

$$\frac{dx_a}{dt} = u_a - g_a(x_a) \qquad \forall a \in \mathcal{A} \tag{9.1}$$

where $\mathcal{A}$ denotes the set of all arcs of the network of interest. Although (9.1) is a fairly obvious identity, it seems to have been first studied in depth by Merchant and Nemhauser (1978a, 1978b) in the context of system optimal dynamic traffic assignment. The same dynamics were employed by Friesz et al. (1989) and Wie et al. (1995) to explore certain extensions of the Merchant-Nemhauser model. Exit flow

functions have been widely criticized as difficult to specify and measure. Exit flow functions are known to allow certain anomalies as illustrated and discussed by Carey (1986, 1987, 1992, 1995). As a consequence, many researchers studying dynamic network flow problems have abandoned dynamics based on exit flow functions.

9.2.2 Dynamics with Controlled Entrance and Exit Flows

A possible modification of the Merchant-Nemhauser arc dynamics that avoids the use of problematic exit flow functions is to treat both arc entrance and exit flows as control variables. Let $\mathcal{W}$ be the set of origin-destination pairs and recall $\mathcal{A}$ is the set of arcs for the network of interest. Then, one way to operationalize the idea of modeling entrance and exit flows as controls is to write

$$\frac{dx_a^{ij}}{dt} = u_a^{ij} - v_a^{ij} \qquad \forall a \in \mathcal{A}, (i,j) \in \mathcal{W} \tag{9.2}$$

where x_a^{ij} is the volume on arc a traveling between origin-destination pair (i,j), while u_a^{ij} and v_a^{ij} denote the rates at which traffic, also traveling between (i,j), enters and exits arc a, respectively. By treating both u_a^{ij} and v_a^{ij} as control variables, we do not mean to imply that any kind of normative considerations have been introduced, for these variables are viewed as controlled by network users constrained by physical reality and observed only at the level of their aggregate (flow) behavior. A criticism is that missing from the unembellished version of (9.2) is any explanation of the queue discipline for the origin-destination flows on the same arc: just as with dynamics based on exit flow functions, we have no way of ensuring that the FIFO queue discipline is enforced without additional constraints or assumptions. Furthermore, use of dynamics (9.2) without additional constraints may result in flow propagation speeds faster than would occur under free flow with no congestion, a rather profound violation of physical reality.

To overcome the difficulties mentioned above, Bernstein et al. (1993), Ran et al. (1993), and Ran and Boyce (1996) have suggested flow propagation constraints for dynamics (9.2) of the form:

$$U_a^p(t) = V_a^p[t + \Delta_a(t)] \qquad \forall a \in \mathcal{A}, p \in \mathcal{P} \tag{9.3}$$

where $U_a^p(.)$ and $V_a^p(.)$ are the cumulative numbers of vehicles associated with path p that are entering and leaving link a, respectively, while $\Delta_a(t)$ denotes the time needed to traverse link a at time t and $\mathcal{P}$ is the set of all paths. The meaning of these constraints is fairly intuitive: vehicles entering an arc at a given moment in time must exit at a later time consistent with the arc traversal time. Moreover, these constraints assume that flows moving through the network are incompressible; that is, wave packets and vehicle platoons are neither shortened nor elongated in the presence of congestion. We will see later that this incompressibility assumption is

incompatible with at least one model of arc delay widely employed in dynamic traffic assignment modeling; this is because the constraints (9.3) omit a fundamental term that describes the expansion and contraction of wave packets or platoons moving through the network.

9.2.3 Cell Transmission Dynamics

The *cell transmission model* is the name given by Daganzo (1994) to dynamics of the following form:

$$x_j(t+1) - x_j(t) = y_j(t) - y_{j+1}(t) \tag{9.4}$$

$$y_j(t) = \min\left\{x_{j-1}(t), Q_j(t), \alpha\left[N_j(t) - x_j(t)\right]\right\} \tag{9.5}$$

where t is now a discrete time index and a unit time step is employed. In the above, the subscript $j \in C$ refers to a spatially discrete physical "cell" of the roadway segement of interest, while $(j-1) \in C$ refers to the cell downstream; C is of course the set of cells needed to describe the roadway. Also, x_j refers to the traffic volume of cell j. Furthermore, y_j is the actual inflow to cell j, Q_j is the maximal rate of discharge from cell j, N_j is the holding capacity of cell j, and α is a parameter. Daganzo (1995) shows how (9.4) and (9.5) can be extended to deal with network structures through straightforward bookkeeping. Note that (9.5) is a constraint on the variables x_j and y_j.

The language introduced previously is readily applicable to the cell transmission model; in particular, (9.4) are arc (cell) dynamics (now several dummy arcs can make up a real physical arc), and (9.5) are flow propagation constraints. The cell transmission model also includes an implicit notion of arc delay. That notion, however, is somewhat subtle: namely delay is that which occurs from traffic flowing in accordance with the fundamental diagram of road traffic. This is because (9.5), as explained by Daganzo (1994), is really a piecewise linear approximation compatible with hydrodynamic models of traffic flow. This feature immunizes the cell transmission model against potential inconsistencies among the three submodels: arc dynamics, flow propagation, and arc delay.

It is possible to couple the dynamical description (4) and (5) to path and departure time choice mechanisms to yield a mathematically exact model for network equilibrium. In fact, Ziliaskopoulos (2000) and Li et al. (2003) have employed the cell transmission model to determine dynamic system optimal flows. Lo and Szeto (2002) and Szeto and Lo (2004) have used the cell transmission model to investigate dynamic user equilibrium. A major difficulty associated with using the cell transmission model as a foundation for a dynamic network user equilibrium model is the fact that the righthand sides of (9.4) are nondifferentiable; this means that, if the path delay operators are nonlinear, any kind of direct control-theoretic approach will involve a nonsmooth Hamiltonian and all the attendant difficulties.

9.2.4 Dynamics Based on Arc Exit Time Functions

Another alternative to the Merchant-Nemhauser dynamics (9.1) is based on the use of *exit time functions and their inverses*. This approach, due to Friesz et al. (1993), allows one to avoid use of exit flow functions and the pitfalls associated therewith. The resulting formulation of link dynamics and of the dynamic network user equilibrium problem has been employed by Wu et al. (1998), Zhu and Marcotte (2000), and Bliemer and Bovy (2003) for additional investigations of dynamic network flows. The main numerical complication of the Friesz et al. (1993) formulation arises form the need to numerically determine inverse exit time functions. Even so, this challenge is not insummountable, as we shall see.

To understand the exit time function, let t_i be the time at which flow exits the ith arc of some path p when departure from the origin of that path has occurred at time t. The relationship of these two instants of time is expressed as

$$t_i = \tau_{a_i}^p (t) \tag{9.6}$$

and we call $\tau_{a_i}^p (.)$ the exit time function for arc a_i of path p. The inverse of the exit time function is written as

$$t = \theta_{a_i}^p (t_i) \tag{9.7}$$

and describes the time of departure t from the origin of path p for flow exiting arc a_i of that path at time t_i. Consequently, the following identity must hold:

$$t = \theta_{a_i}^p \left(\tau_{a_i}^p (t)\right) \tag{9.8}$$

for all time t for which flow behavior is being modeled. The role of the exit time function becomes clearer if we describe path p as the following sequence of conveniently labeled arcs:

$$p \equiv \left\{a_1, a_2, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_{m(p)}\right\} \tag{9.9}$$

where $m(p)$ is the number of arcs in path p. It then follows immediately that the total traversal time for path p may be articulated in terms of the final exit time function and the departure time:

$$D_p (t) = \sum_{i=1}^{m(p)} \left[\tau_{a_i}^p (t) - \tau_{a_i-1}^p (t)\right] = \tau_{a_{m(p)}}^p (t) - t \tag{9.10}$$

when departure from the origin of path p is at time t.

Construction of the arc dynamics begins by noting that arc volumes are the sum of volumes associated with individual paths using a given arc:

$$x_a (t) = \sum_p \delta_{ap} x_a^p (t) \qquad \forall a \in \mathcal{A} \tag{9.11}$$

where $x_a^p(t)$ denotes the volume on arc a associated with path p and

$$\delta_{ap} = \begin{cases} 1 & \text{if arc } a \text{ belongs to path } p \\ 0 & \text{otherwise} \end{cases} \tag{9.12}$$

If we use the notation $h_p(t)$ for the flow entering path p at time t, it is possible to express its contribution to the flow on any arc at a subsequent instant in time using the inverse exit time functions defined previously. In particular, the cumulative departures from the origin for a given path up to some moment t, which is the cumulative number of vehicles that have entered the first arc of that path, may be expressed as

$$I_p(t) = \int_0^t h_p(y)\, dy \qquad \forall p \in \mathcal{P} \tag{9.13}$$

where y is a dummy variable of integration and $\mathcal{P}$ is the set of all paths of the network. From the definition of $I_p(t)$, the volume contributed by a path to any arc at any moment in time is easily represented as the difference between the cumulative departures from the origin that have had time to reach the arc and the cumulative departures from the origin that have had time to exit the arc, or

$$x_{a_i}^p(t) = I_p\left[\theta_{a_{i-1}}^p(t)\right] - I_p\left[\theta_{a_i}^p(t)\right] \qquad \forall a \in \mathcal{A},\ p \in \mathcal{P} \tag{9.14}$$

which becomes

$$x_{a_i}^p(t) = \int_0^{\theta_{a_{i-1}}^p(t)} h_p(y)\, dy - \int_0^{\theta_{a_i}^p(t)} h_p(y)\, dy \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] \tag{9.15}$$

Expressions (9.14) and (9.15) are predicated on the elementary notion that the flow entering arc a_i is the flow exiting its predecessor arc a_{i-1} for any path and any instant in time. It is important to realize that (9.14) and (9.15) are fundamental identities that must apply to any dynamic traffic network. The strongest assumption made in their articulation is that the inverse exit time functions exist.

Note that (9.14) and (9.15) can be expressed in the Merchant-Nemhauser form as

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] \tag{9.16}$$

where the following definitions of entrance and exit flows related to path p are employed:

$$g_{a_{i-1}}^p(t) = \frac{dI_p\left[\theta_{a_{i-1}}^p(t)\right]}{dt} = \frac{d}{dt}\int_0^{\theta_{a_{i-1}}^p(t)} h_p(y)\, dy \tag{9.17}$$

$$g_{a_i}^p(t) = \frac{dI_p\left[\theta_{a_i}^p(t)\right]}{dt} = \frac{d}{dt}\int_0^{\theta_{a_i}^p(t)} h_p(y)\, dy \tag{9.18}$$

These last two expressions can be considerably simplified by using the following identity based on the chain rule and valid for arbitrary j:

$$\frac{d}{dt}\int_0^{\theta^p_{a_j}(t)} h_p(y)\,dy = \frac{d\theta^p_{a_j}(t)}{dt}\cdot\frac{d}{d\theta^p_{a_j}(t)}\int_0^{\theta^p_{a_j}(t)} h_p(y) = \frac{d\theta^p_{a_j}(t)}{dt}\cdot h_p\left[\theta^p_{a_j}(t)\right] \tag{9.19}$$

Use of (9.19) allows (9.17) and (9.18) to be re-expressed as

$$g^p_{a_{i-1}}(t) \equiv \frac{d\theta^p_{a_{i-1}}(t)}{dt}\cdot h_p\left[\theta^p_{a_{i-1}}(t)\right] \tag{9.20}$$

$$g^p_{a_i}(t) \equiv \frac{d\theta^p_{a_i}(t)}{dt}\cdot h_p\left[\theta^p_{a_i}(t)\right] \tag{9.21}$$

Note that even though (9.16) is remarkably similar to (9.2), the entrance and exit flows (9.20) and (9.21) have been very carefully related to departure rates (i.e., path flows) to avoid internal inconsistencies and flow propagation anomalies like instantaneous propagation. Note also that the dynamics (9.16) are intrinsically complicated, having righthand sides that are neither explicit functions nor variables but rather operators that involve inverse exit time functions. Our reading of the literature indicates that relationships (9.20) and (9.21) were first noted by Tobin (1993) and Friesz et al. (1995).

9.2.5 *Constrained Dynamics Based on Proper Flow Propagation Constraints*

There is a way of re-expressing the model of arc dynamics (9.16) to obtain an alternative formulation involving constrained differential equations, state-dependent time lags, and arc entrance and exit flows that are control variables rather than operators. We will see that this alternative formulation obviates the need to explicitly know exit time functions and their inverses but nonetheless preserves all the main features of the Friesz et al. (1993) model of link dynamics. Moreover, the resulting dynamical description may be readily employed as a foundation for a dynamic network user equilibrium model. The constrained dynamical formulation rests on using the definition of $\theta^p_{a_i}(t)$ to rewrite (9.21) as

$$g^p_{a_i}\left[\tau^p_{a_i}(t)\right] = h_p(t)\,\frac{d\theta^p_{a_i}\left[\tau^p_{a_i}(t)\right]}{d\,\tau^p_{a_i}(t)}\,\frac{d\,\tau^p_{a_i}(t)}{dt} \tag{9.22}$$

Furthermore, use of the chain rule and the identity $\theta^p_{a_i}\left[\tau^p_{a_i}(t)\right] = t$ easily reveals that

$$\frac{d\theta^p_{a_i}\left[\tau^p_{a_i}(t)\right]}{d\,\tau^p_{a_i}(t)} = \left[\frac{d\,\tau^p_{a_i}(t)}{dt}\right]^{-1} \tag{9.23}$$

By substituting (9.23) into (9.22), we obtain

$$h_p(t) = g_{a_i}^p\left[\tau_{a_i}^p(t)\right]\frac{d\tau_{a_i}^p(t)}{dt} \tag{9.24}$$

which of course holds for any path $p \in \mathcal{P}$ and any arc $a_i \in p$.

We will next need to model link delay. To this end we introduce a simple deterministic link delay model suggested by Friesz et al. (1993) for modeling dynamic user equilibria and herein named the *separable arc delay model*. To articulate this delay model, let the time to traverse arc a_i for drivers who arrive at its tail node at time t be denoted by $D_{a_i}\left[x_{a_i}(t)\right]$. That is, the time to traverse arc a_i is only a function of the number of vehicles in front of the entering vehicle at the time of entry. As a consequence, we have

$$\tau_{a_1}^p = t + D_{a_1}\left[x_{a_1}(t)\right] \qquad \forall\ p \in \mathcal{P} \tag{9.25}$$

$$\tau_{a_i}^p = \tau_{a_{i-1}}^p(t) + D_{a_i}\left[x_{a_i}\left(\tau_{a_{i-1}}^p(t)\right)\right] \qquad \forall\ p \in \mathcal{P},\ i \in [2, m(p)] \tag{9.26}$$

Differentiating (9.25) and (9.26) with respect to time gives

$$\frac{d\tau_{a_1}^p(t)}{dt} = 1 + D'_{a_1}\left[x_{a_1}(t)\right]\frac{dx_{a_1}(t)}{dt} \qquad \forall\ p \in \mathcal{P} \tag{9.27}$$

$$\frac{d\tau_{a_i}^p(t)}{dt} = \left[1 + D'_{a_i}\left[x_{a_i}\left(\tau_{a_{i-1}}^p(t)\right)\right]\frac{dx_{a_1}\left[\tau_{a_{i-1}}^p(t)\right]}{d\tau_{a_{i-1}}^p(t)}\right]\frac{d\tau_{a_{i-1}}^p(t)}{dt} \qquad \forall\ p \in \mathcal{P},\ i \in [2, m(p)] \tag{9.28}$$

where we have again used the chain rule and the "$\prime$" superscript denotes differentiation with respect to the associated function argument. Thus, we are clearly assuming that all arc delay functions are differentiable with respect to their own arguments, an assumption maintained throughout this chapter.

Evidently, expressions (9.24), (9.25), and (9.27) are easily combined to yield

$$g_{a_1}\left(t + D_{a_1}\left[x_{a_1}(t)\right]\right)\left(1 + D'_{a_1}\left[x_{a_1}(t)\right]\dot{x}_{a_1}(t)\right) = h_p(t) \tag{9.29}$$

where the overdot "$\cdot$" refers to a total time derivative. Proceeding inductively from this last result with the guidance of (9.28), we obtain

$$g_{a_i}^p\left(t + D_{a_i}\left[x_{a_i}(t)\right]\right)\left(1 + D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}(t)\right) = g_{a_{i-1}}^p(t) \qquad \forall\ p \in \mathcal{P},\ i \in [2, m(p)] \tag{9.30}$$

Expressions (9.29) and (9.30) are *flow propagation constraints* derived in a fashion that makes them completely consistent with the chosen exit time function dynamics and the separable arc delay model. Note that these constraints involve a state-dependent time lag $D_{a_i}\left[x_{a_i}(t)\right]$ but make no explicit reference to the exit

time functions and their inverses. Expressions (9.29) and (9.30) may be interpreted as describing the expansion and contraction of vehicle platoons or wave packets moving through various levels of congestion enroute to their destinations. These flow propagation constraints were first pointed out by Tobin (1993) and presented by Friesz et al. (1995). Our reading of the literature indicates Astarita (1995, 1996) independently proposed flow propagation constraints that are essentially identical to (9.29) and (9.30).

To support our development of a dynamic network user equilibrium model in subsequent sections, we need to introduce some additional categories of constraints. The first of these is the flow conservation constraints, which we express as

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \qquad \forall (i,j) \in \mathcal{W} \tag{9.31}$$

where $\mathcal{P}_{ij}$ is the set of all paths that connect origin-destination pair (i,j), $\mathcal{W}$ is set of all origin-destination pairs, t_0 is the initial time, and t_f is the terminal time. Furthermore, Q_{ij} is the fixed travel demand for origin-destination pair (i,j). The fixed travel demand vector is

$$Q = \left(Q_{ij} : (i,j) \in \mathcal{W}\right) \tag{9.32}$$

Note that

$$Q_{ij} = 0 \text{ if } (i,j) \notin \mathcal{W}$$

Finally, we impose the nonnegativity restrictions

$$x \geq 0 \qquad g \geq 0 \qquad h \geq 0 \tag{9.33}$$

where

$$x \equiv \left(x_{a_i}^p : p \in [1, |\mathcal{P}|], i \in [1, m(p)]\right) \tag{9.34}$$

$$g \equiv \left(g_{a_i}^p : p \in [1, |\mathcal{P}|], i \in [1, m(p)]\right) \tag{9.35}$$

$$h \equiv \left(h_p : p \in [1, |\mathcal{P}|]\right) \tag{9.36}$$

are the relevant vectors of state variables and control variables.

As a consequence of the preceding development we can now state the model of constrained arc dynamics we will subsequently employ in modeling user equilibrium:

$$\frac{dx_{a_1}^p(t)}{dt} = h_p(t) - g_{a_1}^p(t) \qquad \forall p \in \mathcal{P} \tag{9.37}$$

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall\, p \in \mathcal{P}, i \in [2, m(p)] \tag{9.38}$$

$$\sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\,dt = Q_{ij} \qquad \forall\,(i,j)\in\mathcal{W} \tag{9.39}$$

$$h_p(t) = g_{a_1}\left(t + D_{a_1}\left[x_{a_1}(t)\right]\right)\left(1 + D'_{a_1}\left[x_{a_1}(t)\right]\dot{x}_{a_1}(t)\right) \tag{9.40}$$

$$g^p_{a_{i-1}}(t) = g^p_{a_i}\left(t + D_{a_i}\left[x_{a_i}(t)\right]\right)\left(1 + D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}(t)\right) \qquad \forall\ p\in\mathcal{P},\ i\in[2,m(p)] \tag{9.41}$$

$$x(t_0) = x_0 \tag{9.42}$$

It should be clear that the link volumes $x^p_{a_i}$ are natural state variables while the path flows h_p and link exit (entrance) flows $g^p_{a_i}$ are natural control variables in the above constrained dynamical formulation. The essential feature of the preceding discussion is: in order to use arc inflows and outflows as control variables one must use flow propagation constraints that are fully consistent with the dynamics selected and the delay model employed. Otherwise one will obtain an intrinsically inconsistent model. We have derived flow propagation constraints for the case of dynamics (9.37) and (9.38) and for delays based on the separable arc delay model. Use of our flow propagation constraints with other dynamics and other delay models would be ill advised.

9.3 The Measure-Theoretic Nature of DUE

In our remarks immediately above, we have assumed large numbers of vehicles and the intrinsic continuity of time allow traffic volumes to be represented by a continuous-time, continuous-state model. However, real-world departure rates are discontinuous in time. This may easily be understood by considering a moment of time when one vehicle enters the first arc of a path; at that instant there is a discontinuous change in the departure rate. This means that, in order to construct a continuous-time and continuous-state model, departure rates should be considered densities that are equivalent so long as they differ only on a set of measure zero and that Lebesgue concepts of integration should be employed, resulting in equilibrium conditions that need only hold almost everywhere.

For convenience, let $\mathcal{L}_+ = (L^2_+[t_0,t_f])^\pi$ denote the nonnegative cone of the π-fold product of the Hilbert space $L^2\left[t_0,t_f\right]$ of square-integrable functions on the closed interval $\left[t_0,t_f\right]$, where

$$\pi = |\mathcal{P}|$$

Each element, $h = (h_p : p\in\mathcal{P})\in\mathcal{L}_+$ is interpreted as a vector of *departure-time densities*, or more simply path flows, measured at the entrance of the first arc of the relevant path. It will be seen that these departure time densities are defined only up to a set of measure zero. With this in mind, let ν denote a Lebesgue measure on

$[t_0, t_f]$, and for each measurable set, $S \subseteq [t_0, t_f]$, let $\forall_\nu(t \in S)$ denote the phrase *for ν-almost all $t \in S$*. If $S = [t_0, t_f]$, then we may at times simply write $\forall_\nu(t)$. Using the notation and concepts we have mentioned, the feasible region for DUE when effective delay operators are known is

$$\Lambda = \left\{ h \in \mathcal{L}_+ : \sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, d\nu(t) = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \right\} \tag{9.43}$$

For simplicity of notation, we shall express the integrals in (9.43) without making explicit reference to ν when measure-theoretic arguments are not needed and confusion will not result.

In light of the observations made above, we may now construct a mathematical statement of DUE as follows. In order to define an appropriate concept of *minimum travel costs* in the present context, we employ the measure-theoretic analogue of the infimum of a set of numbers. In particular, for any measurable set $S \subseteq [t_0, t_f]$ with $\nu(S) > 0$, and any measurable function $M : S \to \Re^1$, the *essential infimum* of some M on S is given by

$$\text{ess inf}\{M(s) : s \in S\} = \sup\{x \in \Re^1 : \nu\{s \in S : M(s) < x\} = 0\} \tag{9.44}$$

Note that, for each $\alpha > \text{ess inf}\{M(s) : s \in S\}$, it must be true by definition that $\nu\{s \in S : M(s) < \alpha\} > 0$. Next, for each $p \in \mathcal{P}$, we define the operator $F_p : [t_0, t_f] \times \mathcal{L}_+ \to \Re^1_+$ for all $(t, h) \in [t_0, t_f] \times \mathcal{L}_+$ by

$$F_p(t,h) = D_p(t,h) + \Theta[t + D_p(t,h) - T_A] \geq 0 \tag{9.45}$$

where $\Theta[.]$ is a penalty for early or late arrival relative to the desired arrival time T_A. We interpret $F_p(t,h)$ as the effective delay at time t on path p under travel conditions h. Presently, our only assumption about such costs is that for each $h \in \mathcal{L}_+$ the operator $F_p(\cdot, h) : [t_0, t_f] \to \Re^1_+$ is measurable on $[t_0, t_f]$. Given these concepts, observe that, for any kl-traveler who is currently considering the choice of a departure time for path $p \in \mathcal{P}_{kl}$ under travel conditions h, the lower bound on *achievable* cost levels for this traveler is given by the essential infimum of $F_p(\cdot, h)$ over the set of all departure times. Hence, the relevant lower bound on such achievable costs is given by

$$\mu_p(h) = \text{ess inf}\{F_p(t,h) : t \in [t_0, t_f]\} \geq 0 \tag{9.46}$$

Given this lower bound on each path $p \in \mathcal{P}_{kl}$ it then follows (from the finiteness of the path set $\mathcal{P}_{kl}$) that, for any kl-traveler who is currently reconsidering his/her present choice of departure time $t \in [t_0, t_f]$ and path $p \in \mathcal{P}_{kl}$, the relevant lower bound on achievable costs for this traveler is given by

$$\mu_{kl}(h) = \min\{\mu_p(h) : p \in \mathcal{P}_{kl}\} \geq 0 \tag{9.47}$$

It is important to note that because of the arrival penalty function and congestion externalities of the network, there may be more than one departure time for the which users will incur μ_{kl}. In fact, there may be an interval of departure times for which μ_{kl} is realized.

With these concepts, we are now ready to define the relevant notion of an *equilibrium* for simultaneous path choice and departure time decisions:

Definition 9.1. *Dynamic user equilibrium. For any* $h = (h_p : p \in \mathcal{P}) \in \Lambda$ *and any nonnegative vector* $\mu = (\mu_{kl} : (k,l) \in \mathcal{W}) \in \Re_+^{|\mathcal{W}|}$, *the pair* (h, μ) *is a simultaneous departure-time-and-path-choice dynamic user equilibrium if and only if the following two conditions are satisfied for all* $p \in \mathcal{P}_{kl}$ *and for all* $(k,l) \in \mathcal{W}$*:*

$$h_p(t) > 0 \Rightarrow F_p(t,h) = \mu_{kl} \quad \forall_\nu(t) \tag{9.48}$$

$$F_p(t,h) \geq \mu_{kl} \quad \forall_\nu(t) \tag{9.49}$$

To interpret these conditions, observe that, since $h_p(t) > 0$ must hold on some set of positive measure for at least one $p \in \mathcal{P}_{kl}$, it follows from (9.48) that the μ_{kl} are precisely the essential infima of cost levels achievable by kl-travelers at all available departure times under h. Given these observations, condition (9.48) is seen to assert that every traveler in the system is currently achieving a cost level that cannot be improved by changing his/her current choice of path and/or departure time. Furthermore, for almost all t, if $F_p(t,k) > \mu_{kl}$, from (9.48), $h_p(t) = 0$. Thus, the dynamic equilibrium conditions (9.48) and (9.49) are directly analogous to the usual static conditions for user optimized flow, requiring that costs be minimal for the current path and departure time choices for ν-almost all kl-travelers.

Observe that the path flows h_p may be viewed as densities that are only unique up to sets of ν-measure zero. (Formally, they are equivalent classes of functions differing only on sets of ν-measure zero.) Hence, although the notion of a single "kl-traveler" serves as a convenient story for purposes of behavioral interpretation, such individuals are formally sets of ν-measure zero, and can have no influence on the densities h_p. Thus the only meaningful notion of "dynamic equilibrium" here is one in which no set of kl-travelers of positive measure can all do better by changing their current decisions. However, the lower bound on costs which are achievable by switching sets of kl-travelers of positive measure under h is precisely the essential infimum, $\mu_{kl}(h)$ as defined by (9.46) and (9.47). Thus, these costs are the appropriate ones for defining user equilibria in a dynamic setting.

9.4 The Infinite-Dimensional Variational Inequality Formulation

Friesz et al. (1993) observe that the integral equation form of the arc dynamics (9.15) may be used to eliminate the state variables and arc exit flows completely from the formulation of DUE, for then the effective path delay operators are expressible

solely in terms of departure rates and the time of departure; that is, the effective delay operators are of the form

$$F_p(t,h) \quad \forall p \in \mathcal{P} \tag{9.50}$$

The operators (9.50) can also be obtained from a simulation or response service methodology for a particular network of interest. Friesz et al. (1993) show that any solution of the following variational inequality is also a solution of the DUE problem:

$$\left.\begin{array}{c} h^* \in \Lambda \\ \langle F(t,h^*), (h-h^*)\rangle \geq 0 \quad \forall h \in \Lambda \end{array}\right\} \tag{9.51}$$

where

$$\left\langle F\left(t,h^*\right), \left(h-h^*\right)\right\rangle \equiv \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} F_p\left(t,h^*\right)\left[h_p(t) - h_p^*(t)\right] dt,$$

and we use the feasible region introduced previously:

$$\Lambda = \left\{ h \in \mathcal{L}_+ : \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \right\} \tag{9.52}$$

We refer, in this chapter, to the formulation (9.51) as *VI(F,Λ)*. The advantage of this formulation is that it subsumes almost all DUE models regardless of the arc dynamics, flow propagation constraints and arc delay functions employed; it is the formulation originally put forward by Friesz et al. (1993).

To develop a variational inequality formulation of DUE, we first establish the following elementary property of measurable functions on the real interval $[t_0, t_f]$:

Lemma 9.1. *For any set $S \subseteq [t_0, t_f]$ with $\nu(S) > 0$ and any measurable function with $\nu\{t \in S : f(t) > 0\} > 0$, there is some $\epsilon_0 > 0$ such that $\nu\{t \in S : f(t) > \epsilon\} > 0$ for all $\epsilon \in [0, \epsilon_0]$.*

Proof. If for each $n > 0$ we set

$$S_n = \{t \in S : 1/n < f(t)\},$$

then by definition

$$\{t \in S : f(t) > 0\} = \bigcup_n S_n$$

However, by the countable subadditivity of measures (Halmos, 1974, Theorem 8.*C*),

$$0 < \nu\{t \in S : f(t) > 0\} = \nu\left(\bigcup_n S_n\right) \leq \sum_n \nu(S_n)$$

which in turn implies that $\nu(S_n) > 0$ for some $n > 0$. Hence, by letting $\epsilon_0 = 1/n > 0$, we may conclude that for all $\epsilon \in [0, \epsilon_0]$

$$S_n \subseteq \{t \in S : f(t) > \epsilon\} \Rightarrow \nu\{t \in S : f(t)\} > 0\} > 0$$

■

Given Lemma 9.1, we are ready to establish the following equivalent formulation of simultaneous path and departure equilibria:

Theorem 9.1. *Infinite-dimensional inequality formulation of DUE. The simultaneous departure-time-and-path-choice dynamic user equilibrium of Definition 9.1 is equivalent to the following variational inequality problem on Λ: find $h^* \in \Lambda$ such that for all $h \in \Lambda$:*

$$\sum_{p \in \mathcal{P}} \int_{t_0}^{t_f} F_p(t, h^*)[h_p(t) - h_p^*(t)] d\nu(t) \geq 0 \tag{9.53}$$

We refer to this formulation as VI(F, Λ).

Proof. We repeat here the proof by Friesz et al. (1993). That proof is in two parts.

(i)[*Necessity*] If (h^*, μ^*) is a dynamic user equilibrium then $h^* \in \Lambda$ by definition. Hence, to establish that h^* is a solution to the (9.53), it suffices to show that for all $h \in \Lambda$ and $(k, l) \in \mathcal{W}$

$$\sum_{p \in \mathcal{P}_{kl}} \int_{t_0}^{t_f} F_p(t, h^*)[h_p(t) - h_p^*(t)] d\nu(t) \geq 0 \tag{9.54}$$

However, because (9.52) implies that

$$\int_{t_0}^{t_f} [h_p(t) - h_p^*(t)] d\nu(t) = 0$$

it follows that (9.54) is equivalent to the condition that

$$\sum_{p \in \mathcal{P}_{kl}} \int_{t_0}^{t_f} \{F_p(t, h^*) - \mu_{kl}^*\}[h_p(t) - h_p^*(t)] d\nu(t) \geq 0 \quad (k, l) \in \mathcal{W} \tag{9.55}$$

Hence, it suffices to show that for all $h \in \Lambda$, $p \in \mathcal{P}_{kl}$ and $(k, l) \in \mathcal{W}$

$$\{F_p(t, h^*) - \mu_{kl}^*\}[h_p(t) - h_p^*(t)] \geq 0, \quad \forall_\nu(t) \tag{9.56}$$

To do so, observe first that if (9.56) fails for any $t \in [t_0, t_f]$, then either $F_p(t, h^*) - \mu_{kl}^* < 0$ or $h_p(t) - h_p^*(t) < 0$. But by (9.49), for ν-almost all t, $F_p(t, h^*) - \mu_p^* \geq 0$. Moreover, by (9.48) it follows that for ν-almost all t

$$h_p(t) < h_p^*(t) \Rightarrow h_p^*(t) > 0 \Rightarrow F_p(t, h^*) = \mu_{kl}^* \\ \Rightarrow \{F_p(t, h^*) - \mu_{kl}^*\}[h_p(t) - h_p^*(t)] = 0 \tag{9.57}$$

Hence (9.56) follows.

(ii) [*Sufficiency*] Next suppose that $h^* \in \Lambda$ satisfies (9.53) for all $h \in \Lambda$, and let the individual components of the vector $\mu^* = (\mu_{kl}^* : (k,l) \in \mathcal{W}) \in \Re_+^{|\mathcal{W}|}$ be defined by $\mu_{kl}^* = \mu_{kl}(h^*)$ for all $(k,l) \in \mathcal{W}$. Our objective is to show that (h^*, μ^*) is a dynamic user equilibrium. To do so, observe first from the definitions of $\mu_p(h^*)$ and $\mu_{kl}(h^*)$ above that, for all $p \in \mathcal{P}_{kl}$ and ν-almost all t, that $F_p(t, h^*) \geq \mu_p(h^*) \geq \mu_{kl}(h^*) = \mu_{kl}^*$. Hence, (h^*, μ^*) satisfies condition (9.49) by construction, and it remains to establish condition (9.48). To do so, suppose to the contrary that (9.48) fails for some $p \in \mathcal{P}_{kl}$, $(k,l) \in \mathcal{W}$. Then, by definition, the set

$$S_p = \{t \in [t_0, t_f] : h_p^*(t) > 0, F_p(t, h^*) - \mu_{kl}^* > 0\} \tag{9.58}$$

must have positive measure. In particular, this implies from Lemma 9.1 that for some sufficiently small value of $\epsilon > 0$ the subset

$$S_p(\epsilon) = \{t \in S_p : F_p(t, h^*) - \mu_{kl}^* > 2\epsilon\} \tag{9.59}$$

has positive measure. Since $S_p(\epsilon) \subseteq S_p \Rightarrow h_p^*(t) > 0$, $\forall_\nu[t \in S_p(\epsilon)]$, a second application of Lemma 9.1 shows that there exists some sufficiently small value of $\delta > 0$ such that

$$S_p(\epsilon, \delta) = \{t \in S_p(\epsilon) : h_p^*(t) > \delta\} \tag{9.60}$$

has positive measure. Next, choosing any path $q \in \mathcal{P}_{kl}$ with $\mu_q(h^*) = \mu_{kl}(h^*)$ (possibly with $q = p$), it follows from the definition of $\mu_p(h^*)$ that the set

$$T_q(\epsilon) = \{t \in [t_0, t_f] : C_q(t, h^*) < \mu_{kl}^* + \epsilon\} \tag{9.61}$$

also has positive measure. Finally, letting $\alpha_0 = \min\{\nu[S_p(\epsilon, \delta)], \nu[T_q(\epsilon)]\} > 0$ and observing that the Lebesgue measure is nonatomic, it follows (Halmos, 1974, Proposition 41.2) that for any choice of $\alpha \in (0, \alpha_0)$ there exist subsets $S_p(\epsilon, \delta, \alpha) \subseteq S_p(\epsilon, \delta)$ and $T_q(\epsilon, \alpha) \subseteq T_q(\epsilon)$ with $\nu[S_p(\epsilon, \delta, \alpha)] = \alpha = \nu[T_q(\epsilon, \alpha)]$. Given these two sets, we now construct a vector of densities $h = (h_r : r \in \mathcal{P}) \in \Lambda$ which violates condition (9.53) for h^*. To do so, let $h_r = h_r^*$ for all $r \in \mathcal{P} - \{p, q\}$, and let h_p and h_q be defined respectively by

$$h_p(t) = \begin{cases} h_p^*(t) - \delta & t \in S_p(\epsilon, \delta, \alpha) \\ h_p^* & t \in [0, T] - S_p(\epsilon, \delta, \alpha) \end{cases} \tag{9.62}$$

$$h_q(t) = \begin{cases} h_q^* + \delta & t \in T_q(\epsilon, \alpha) \\ h_q^*(t) & t \in [0, T] - T_q(\epsilon, \alpha) \end{cases} \tag{9.63}$$

Note that if $p = q$, then these two conditions still yield a well-defined function, h_p. To see this, observe from (9.59) that $S_p(\epsilon, \delta, \alpha) \subseteq S_p(\epsilon)$ implies $F_p(t, h^*) >$

$\mu_{kl}^* + 2\epsilon$, $\forall_\nu[t \in S_p(\epsilon,\delta,\alpha)]$, and similarly from (9.61) that $T_p(\epsilon,\alpha) \subseteq T_p(\epsilon)$ implies $F_p(t,h^*) < \mu_{kl}^* + \epsilon$, $\forall_\nu[t \in T_p(\epsilon,\alpha)]$. Hence, if $p = q$, then we must have $\nu[S_p(\epsilon,\delta,\alpha) \cap T_p(\epsilon,\alpha)] = 0$, and it follows that h_p is well defined up to a set of measure zero. Thus, without loss of generality, we may henceforth assume that $p \neq q$. With this convention, we next show that $h \in \Lambda$. To do so, observe first that for ν-almost all $t \in S_p(\epsilon,\delta,\alpha)$, we have $h_p^*(t) \geq \delta \Rightarrow h_p(t) \geq 0$. Similarly, for ν-almost all $t \in T_q(\epsilon,\alpha)$, $h_q^*(t) \geq 0 \Rightarrow h_q(t) \geq 0$. Moreover,

$$\nu[S_p(\epsilon,\delta,\alpha)] = \alpha = \nu[T_q(\epsilon,\alpha)]$$

implies that

$$\begin{aligned}
\sum_{r\in\mathcal{P}_{kl}} \int_{t_0}^{t_f} h_r(t)d\nu(t) &= \sum_{r\in\mathcal{P}_{kl}-\{p,q\}} \int_{t_0}^{t_f} h_r(t)d\nu(t) \\
&\quad + \int_{t_0}^{t_f} h_p(t)d\nu(t) + \int_{t_0}^{t_f} h_q(t)d\nu(t) \\
&= \sum_{r\in\mathcal{P}_{kl}\setminus\{p,q\}} \int_{t_0}^{t_f} h_r^*(t)d\nu(t) \\
&\quad + \left[\int_{t_0}^{t_f} h_p^*(t)d\nu(t) - \delta\cdot\alpha\right] \\
&\quad + \left[\int_{t_0}^{t_f} h_q^*(t)d\nu(t) + \delta\cdot\alpha\right] \\
&= \sum_{r\in\mathcal{P}_{kl}} \int_{t_0}^{t_f} h_r^*(t)d\nu(t) \\
&= Q_{kl} \qquad (9.64)
\end{aligned}$$

Therefore, we must have $h \in \Lambda$. However, (9.62) and (9.63) also imply that

$$\begin{aligned}
\sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} F_p(t,h^*)[h_p(t) - h_p^*(t)]d\nu(t) &= \int_{S_p(\epsilon,\delta,\alpha)} F_p(t,h_*)[h_p(t) - h_p^*(t)] \\
&\quad + \int_{T_q(\epsilon,\alpha)} C_q(t,h^*)[h_q(t) - h_q^*(t)], \qquad (9.65)
\end{aligned}$$

Furthermore, the construction of $S_p(\epsilon,\delta,\alpha)$ and $T_q(\epsilon,\alpha)$ imply, respectively, that

$$\begin{aligned}
\int_{S_p(\epsilon,\delta,\alpha)} F_p(t,h^*)[h_p(t) - h^*(t)] &= \int_{S_p(\epsilon,\delta,\alpha)} F_p(t,h^*)[-\delta]d\nu(t) \\
&\leq -\int_{S_p(\epsilon,\delta,\alpha)} \left[(\mu_{kl}^* + 2\epsilon)\delta\right] d\nu(t) \\
&= -(\mu_{kl}^* + 2\epsilon)\delta\alpha \qquad (9.66)
\end{aligned}$$

and

$$\begin{aligned}\int_{T_q(\epsilon,\alpha)} C_q(t,h^*)[h_q(t)-h_q^*(t)] &= \int_{T_q(\epsilon,\alpha)} C_q(t,h^*)[\delta]d\nu(t)\\ &\le \int_{T_q(\epsilon,\alpha)} \left[(\mu_{kl}^*+\epsilon)\delta\right] d\nu(t)\\ &= (\mu_{kl}^*+\epsilon)\delta\alpha \end{aligned} \tag{9.67}$$

Finally, by combining (9.65),(9.66), and (9.67), we may conclude that

$$\sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} F_p(t,h^*)[h_p - h_p^*(t)]d\nu(t) \le -\epsilon\delta\alpha < 0, \tag{9.68}$$

which contradicts (9.53) for this choice of $h \in \Lambda$. Thus the hypothesized failure of condition (9.48) leads to a contradiction, and we may conclude that (h^*, μ^*) is a simultaneous path-departure equilibrium. ■

In Chapter 5 we learned that under certain conditions a finite-dimensional variational inequality may be viewed as a necessary condition for an appropriately defined finite-dimensional extremal problem. In static traffic assignment, the existence of extremal problems corresponding to variational inequality formulations for static user equilibrium requires that the cost (or delay) functions have a symmetric Jacobian matrix. [See Friesz et al. (1985) for a review of static network equilibrium.] The extremal problem is then the minimization of the sum of integrals of arc costs (or delays) when travel demand is exogenous. In the present dynamic case, the usual symmetry conditions ensuring that a line integral has a value independent of the path of integration are not readily tested since the $F_p(t,h)$ operators do not have a closed-form representation. Furthermore, intuitive arguments with regard to the irreversibility of time would seem to contravene such symmetry, for otherwise travelers' decisions at opposite ends of the analysis time horizon would have equal and symmetric impacts on one another.

9.5 When Delays Are Exogenous

Let us consider a circumstance for which the effective path delay operators are known in advance or are represented by a simulation model that does not employ the constrained state dynamics (9.37), (9.38), (9.39), (9.40), (9.41), and (9.42). We begin, by noting that the flow conservation constraints

$$\sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\,dt = Q_{ij} \tag{9.69}$$

of formulation (9.51) may be restated as

$$\frac{dy_{ij}}{dt} = \sum_{p \in \mathcal{P}_{ij}} h_p(t) \quad \forall (i,j) \in \mathcal{W} \tag{9.70}$$

$$y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \tag{9.71}$$

$$y_{ij}(t_f) = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \tag{9.72}$$

For discussions in subsequent sections, it is useful to restate the state dynamics (9.70), (9.71), and (9.72) as

$$\frac{dy}{dt} = Ah \tag{9.73}$$

$$y(t_0) = y_0 \equiv 0 \tag{9.74}$$

$$\Psi\left[y(t_f)\right] \equiv Q - y(t_f) = 0 \tag{9.75}$$

where

$$A = \left(A_{ij}^p : (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}\right) \tag{9.76}$$

is the path-OD incidence matrix and

$$A_{ij}^p = \begin{cases} 1 & if \quad p \in \mathcal{P}_{ij} \\ 0 & if \quad p \notin \mathcal{P}_{ij} \end{cases}$$

Also

$$y = \left(y_{ij} : (i,j) \in \mathcal{W}\right)$$

is the vector of dummy variables used to represent the flow conservation constraints. Clearly (9.73), (9.74), and (9.75) constitute a two-point boundary-value problem. We also introduce the inner product notation

$$\langle F^0(t,h^*), h - h^* \rangle = \sum_{p \in \mathcal{P}} \int_{t_0}^{t_f} F_p^0(t,h^*)\left(h_p - h_p^*\right) dt \geq 0 \tag{9.77}$$

where

$$F^0(t,h^*) = \left(F_p^0(t,h^*) : p \in \mathcal{P}\right) \tag{9.78}$$

is the exogenously determined effective path delay operator. As a consequence problem (9.51) may be expressed as the following differential variational inequality: find $h^* \in \Lambda_0$ such that

$$\left.\begin{array}{c} h^* \in \Lambda_0 \\ \langle F^0(t,h^*), h - h^* \rangle \geq 0 \quad \forall h \in \Lambda_0 \end{array}\right\} \quad DVI(F^0, \Lambda_0) \tag{9.79}$$

where we have suppressed the previous measure-theoretic notation, although the integral above remains a Lebesgue integral. Furthermore, in (9.79) the set of admissible controls is now stated as

$$\Lambda_0 = \left\{ h \geq 0 : \frac{dy}{dt} = Ah,\ y\left(t_0\right) = y_0,\ y\left(t_f\right) = Q \right\} \tag{9.80}$$

The variational inequality (9.79) may be written as

$$\sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} F_p^0(t,h^*) h_p dt \geq \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} F_p^0(t,h^*) h_p^* dt \quad \forall h \in \Lambda_0 \tag{9.81}$$

which means that the solution $h^* \in \Lambda_0$ satisfies the optimal control problem

$$\min J_0 = \sum_{(i,j)\in\mathcal{W}} v_{ij} \left[Q_{ij} - y_{ij}\left(t_f\right) \right] + \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} F_p^0(t,h^*) h_p dt \tag{9.82}$$

subject to

$$\frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} h_p(t) \quad \forall (i,j) \in \mathcal{W} \tag{9.83}$$

$$y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \tag{9.84}$$

$$h \geq 0 \tag{9.85}$$

where the v_{ij} are dual variables for the terminal conditions on the state variables. The Hamiltonian for problem (9.82), (9.83), and (9.85) is

$$H_0 = \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} F_p^0(t,h^*) h_p + \sum_{(i,j)\in\mathcal{W}} \lambda_{ij} \sum_{p\in\mathcal{P}_{ij}} u_p \tag{9.86}$$

where the adjoint equation is

$$\frac{d\lambda_{ij}}{dt} = -\frac{\partial H_0}{\partial y_{ij}} = 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.87}$$

with transversality condition

$$\lambda_{ij}\left(t_f\right) = \frac{\partial \sum_{(i,j)\in\mathcal{W}} v_{ij}\left[Q_{ij} - y_{ij}\left(t_f\right)\right]}{y_{ij}\left(t_f\right)} = -v_{ij} = \text{constant} \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.88}$$

The implication of (9.87) and (9.88) is of course that

$$\lambda_{ij}(t) = -v_{ij} \quad \forall (i,j) \in \mathcal{W}, t \in \left[t_0, t_f\right] \tag{9.89}$$

The minimum principle requires the controls h to obey

$$\min H_0 \quad \text{s.t.} \quad -h \leq 0 \tag{9.90}$$

with Kuhn-Tucker conditions

$$F_p^0(t, h^*) - v_{ij} = \rho_p \geq 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.91}$$

where the ρ_p are dual variables satisfying the complementary slackness conditions

$$\rho_p h_p = 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.92}$$

From (9.91) and (9.92) we have immediately the conditions of a dynamic user equilibrium, namely

$$h_p^* > 0, p \in \mathcal{P}_{ij} \Longrightarrow F_p^0(t, h^*) = v_{ij} \tag{9.93}$$

$$F_p^0(t, h^*) > v_{ij}, p \in \mathcal{P}_{ij} \Longrightarrow h_p^* = 0 \tag{9.94}$$

with the obvious interpretation that each dual variable v_{ij} is the essential infimum of the effective unit path delay $F_p^0(t, h^*)$. Thus, we are assured that any solution of our differential variational inequality is a dynamic user equilibrium relative to path and departure time choice.

It is an easy matter to show that (9.93) and (9.94) lead to an appropriate version of (9.79); in fact, the arguments employed in the proof of Theorem 9.1 directly apply. However, it is instructive to give an informal derivation of the same result for those readers not familar with measure theory. In particular we note that

$$F_p^0(t, h^*) \geq v_{ij} \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.95}$$

Therefore, if $\left(h_p - h_p^*\right) \geq 0$ we have

$$F_p^0(t, h^*)\left(h_p - h_p^*\right) \geq v_{ij}\left(h_p - h_p^*\right) \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.96}$$

However,

$$\left(h_p - h_p^*\right) < 0 \Longrightarrow h_p^* > h_p \geq 0 \Longrightarrow h_p^* > 0 \tag{9.97}$$

which by (9.93) requires (9.95) to hold as an equality, thereby assuring (9.96) is valid for any $h_p, h_p^* \in \Lambda_0$. As a consequence, we may sum and integrate both sides of (9.96) to obtain

$$\sum_{p \in \mathcal{P}} \int_{t_0}^{t_f} F_p^0(t, h^*)\left(h_p - h_p^*\right) dt \geq \sum_{p \in \mathcal{P}} \int_{t_0}^{t_f} v_{ij}\left(h_p - h_p^*\right) \tag{9.98}$$

$$= \sum_{(i,j)\in\mathcal{W}} v_{ij} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \left(h_p - h_p^*\right) \tag{9.99}$$

$$= \sum_{(i,j)\in\mathcal{W}} v_{ij}(Q_{ij} - Q_{ij}) = 0 \tag{9.100}$$

Thus, we have established the following result:

Theorem 9.2. *Variational inequality $VI(F^0, \Lambda)$ with exogenous delay operators and differential variational inequality $DVI(F^0, \Lambda_0)$ are equivalent and their solutions are dynamic user equilibria.*

We now need to say a few things about solving $DVI(F^0, \Lambda_0)$.

In particular, we are interested in re-stating (9.79) as a fixed-point problem, primarily because fixed-point problems enjoy simple iterative algorithms that do not involve derivatives; this will allow us to avoid differentiating the effective path delay operators, which are typically nondifferentiable. Consider the following fixed-point problem:

$$h = P_{\Lambda_0}\left[h - \alpha F^0(t, h^*)\right] \quad \Big\} \quad FPP(F^0, \Lambda_0) \tag{9.101}$$

where $P_{\Lambda_0}[.]$ denotes a minimum norm projection and $\alpha \in \Re^1_{++}$ is an arbitrary scalar parameter that may be adjusted to facilitate convergence. We need to establish that any solution of (9.101) is a solution of (9.79). This is done by noting that (9.101) may be restated as

$$h^* = \arg\left\{\min_z \int_{t_0}^{t_f} \frac{1}{2} \sum_{p\in\mathcal{P}} \left[h_p^* - \alpha F_p^0(t, h^*) - z_p\right]^2 dt \quad \text{s.t.} \quad z \in \Lambda_0\right\} \tag{9.102}$$

That is, (9.102) requires that the following optimal control problem must be solved

$$\min_z \sum_{(i,j)\in\mathcal{W}} \gamma_{ij}\left[Q_{ij} - y_{ij}\left(t_f\right)\right] + \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \frac{1}{2}\left[h_p^* - \alpha F_p^0(t, h^*) - z_p\right]^2 dt \tag{9.103}$$

subject to

$$\frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} z_p(t) \quad \forall (i,j) \in \mathcal{W} \tag{9.104}$$

$$y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \tag{9.105}$$

$$z \geq 0 \tag{9.106}$$

The Hamiltonian for this problem is

$$H_0 = \frac{1}{2} \sum_{p\in\mathcal{P}} \left[h_p^* - \alpha F_p^0(t, h^*) - z_p\right]^2 + \sum_{(i,j)\in\mathcal{W}} \eta_{ij} \sum_{p\in\mathcal{P}_{ij}} z_p \tag{9.107}$$

where the adjoint equation is

$$\frac{d\eta_{ij}}{dt} = -\frac{\partial H_0}{\partial y_{ij}} = 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.108}$$

with transversality condition

$$\eta_{ij}\left(t_f\right) = \frac{\partial \sum_{(i,j)\in\mathcal{W}} \gamma_{ij}\left[Q_{ij} - y_{ij}\left(t_f\right)\right]}{y_{ij}\left(t_f\right)} = -\gamma_{ij} = \text{constant}$$
$$\forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.109}$$

Thus, we conclude

$$\eta_{ij}(t) = -\gamma_{ij} \quad \forall (i,j) \in \mathcal{W}, t \in \left[t_0, t_f\right] \tag{9.110}$$

The minimum principle requires the controls z to obey

$$\min H_0 \text{ s.t. } -z \leq 0 \tag{9.111}$$

for which the Kuhn-Tucker conditions include

$$-h_p + \alpha F_p^0(t, h^*) + z_p + \eta_{ij} = \sigma_p \geq 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.112}$$

where the σ_p are dual variables satisfying the complementary slackness conditions

$$\sigma_p z_p = 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.113}$$

Thus

$$z_p = \sigma_p - \eta_{ij} + h_p^* - \alpha F_p^0(t, h^*) \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.114}$$

Using (9.114) with (9.102) we obtain

$$h_p^* = \sigma_p - \eta_{ij} + h_p^* - \alpha F_p^0(t, h^*) \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.115}$$

In light of (9.110) this last result becomes

$$F_p^0(t, h^*) = \frac{\sigma_p}{\alpha} + \frac{\gamma_{ij}}{\alpha} \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in \left[t_0, t_f\right] \tag{9.116}$$

If we define

$$\rho_p \equiv \frac{\sigma_p}{\alpha} \geq 0 \quad \forall (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij} \tag{9.117}$$

$$v_{ij} \equiv \frac{\gamma_{ij}}{\alpha} \quad \forall (i,j) \in \mathcal{W}, \tag{9.118}$$

then, from the complementary slackness conditions (9.113) and the identity $h^* = z$ obtained from (9.102), we see that (9.116) assures the dynamic user equilibrium conditions

$$h_p^* > 0, p \in \mathcal{P}_{ij} \Longrightarrow F_p^0(t, h^*) = v_{ij} \tag{9.119}$$

$$F_p^0(t, h^*) > v_{ij}, p \in \mathcal{P}_{ij} \Longrightarrow h_p^* = 0 \tag{9.120}$$

Manipulations like those intrinsic to (9.97), (9.98), (9.99), and (9.100) establish that (9.119) and (9.120) lead to $DVI(F^0, \Lambda_0)$. Thus, we have established the following result:

Theorem 9.3. *Fixed-point representation of DUE. The differential variational inequality DVI*(F^0, Λ_0) *and fixed-point problem* $FPP(F^0, \Lambda_0)$ *for exogenous delay operators are equivalent and any solutions of them are dynamic user equilibria.*

Associated with our fixed-point formulation $FPP\left(F^0, \Lambda_0\right)$, as we have noted in previous chapters, is a simple and obvious algorithm:

$$h^{k+1} = P_{\Lambda_0}\left[h^k - \alpha F^0(t, h^k)\right] \tag{9.121}$$

When the projection operator is properly interpreted, algorithm (9.121) involves the repeated solution of an optimal control problem and takes the following form:

Step 0. Initialization. Identify an initial feasible solution $h^0 \in \Lambda_0$ and set $k = 0$.

Step 1. Optimal control subproblem. Solve

$$\min_u \sum_{(i,j)\in\mathcal{W}} \gamma_{ij}\left[Q_{ij} - y_{ij}\left(t_f\right)\right] + \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \frac{1}{2}\left[h_p^k - \alpha F_p^0(t, h^k) - u_p\right]^2 dt \tag{9.122}$$

subject to

$$\frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} u_p(t) \quad \forall (i,j) \in \mathcal{W} \tag{9.123}$$

$$y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \tag{9.124}$$

$$u \geq 0 \tag{9.125}$$

and call the solution h^{k+1}.

Step 2. Stopping test. If

$$\left\| h^{k+1} - h^k \right\| \leq \varepsilon$$

where $\varepsilon \in \Re_{++}^1$ is a preset tolerance, stop and declare $h^* \approx h^{k+1}$. Otherwise set $k = k + 1$ and go to Step 1.

In order to to solve the subproblem (9.122), (9.123), (9.124), and (9.125), a critical step is finding the dual variables γ_{ij} for all $(i, j) \in \mathcal{W}$. If these may be approximated using primal information, a direct solution of the subproblem is possible. We note that the optimality conditions for the subproblem provide the following relationships when the current optimal values of each u_p, σ_p and γ_p are referred to as h_p^{k+1}, σ_p^{k+1} and γ_p^{k+1}, respectively:

$$h_p^{k+1} = h_p^k + \sigma_p^{k+1} + \gamma_{ij}^{k+1} - \alpha F_p^0(t, h^k) \quad \forall (i, j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in [t_0, t_f]$$
$$0 = \sigma_p^{k+1} h_p^{k+1} \quad \sigma_p^{k+1} \geq 0 \quad \forall (i, j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in [t_0, t_f]$$

Therefore

$$h_p^{k+1} = \left[h_p^k + \gamma_{ij}^{k+1} - \alpha F_p^0(t, h^k) \right]_+ \quad \forall (i, j) \in \mathcal{W}, p \in \mathcal{P}_{ij}, t \in [t_0, t_f] \tag{9.126}$$

where $[.]_+$ is the elementary projection operator:

$$[x]_+ = \begin{cases} x & if \quad x \geq 0 \\ 0 & if \quad x < 0 \end{cases} \tag{9.127}$$

By virtue of flow conservation, the dual variables γ_{ij}^k obey the following system of uncoupled equations:

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} \left[h_p^k + \gamma_{ij}^{k+1} - \alpha F_p^0(t, h^k) \right]_+ dt = Q_{ij} \quad \forall (i, j) \in \mathcal{W} \tag{9.128}$$

One dimensional line search may be used to find each dual variable γ_{ij}^{k+1} from (9.128). In fact, there are $|\mathcal{W}|$ such line searches to perform, and all may be carried out simultaneously.

9.6 When the Delay Operators Are Endogenous

In Section 9.4 we studied the problem of finding a dynamic user equilibrium when the delay operators are exogenous. Now we consider the circumstance where the delay operators are endogenous. This means that we must employ as constraints the system of constrained dynamics summarized at the end of Section 9.2.5 and repeated here for convenience:

$$\frac{dx_{a_1}^p(t)}{dt} = h_p(t) - g_{a_1}^p(t) \qquad \forall p \in \mathcal{P} \tag{9.129}$$

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall\, p \in \mathcal{P}, i \in [2, m(p)] \quad (9.130)$$

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \qquad \forall\, (i,j) \in \mathcal{W} \quad (9.131)$$

$$h_p(t) = g_{a_1}\left(t + D_{a_1}\left[x_{a_1}(t)\right]\right)\left(1 + D'_{a_1}\left[x_{a_1}(t)\right]\dot{x}_{a_1}(t)\right) \quad (9.132)$$

$$g_{a_{i-1}}^p(t) = g_{a_i}^p\left(t + D_{a_i}\left[x_{a_i}(t)\right]\right)\left(1 + D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}(t)\right) \qquad \forall\, p \in \mathcal{P},\ i \in [2, m(p)] \quad (9.133)$$

$$x(t_0) = x_0 \quad (9.134)$$

That is, we consider the admissible set

$$\Lambda_1 = \{h : (9.129), (9.130), (9.131), (9.132), (9.133), \text{and } (9.134) \text{ hold}\}$$

In particular we seek $h^* \in \Lambda_1$ such that

$$\langle F(t,h), h - h^* \rangle = \sum_{p \in \mathcal{P}} \int_{t_0}^{t_f} F_p\left(t, x\left(h^*\right)\right)\left[h_p(t) - h_p^*(t)\right] dt \geq 0 \quad \forall h \in \Lambda_1 \quad (9.135)$$

We refer to problem (9.135) as $DVI(F, \Lambda_1)$, where the delay operator is now F to signify that path delays are determined endogenously. Bernstein et al. (1993) postulated a similar formulation but used *intuitive* flow propagation constraints that do not account for the compressibility of vehicle platoons. Bernstein et al. (1993) informally argued that their formulation was correct since it was constructed from valid submodels; that is, they did not formally demonstrate that the associated necessary conditions ensured solutions were dynamic network user equilibria. The specific formulation (9.135) was first conjectured by Friesz et al. (1995), also without a formal analysis of the necessary conditions. The first complete analysis of the necessary conditions for (9.135) was carried out by Friesz et al. (2001). The main motivation for offering formulation (9.135) is that a variational inequality is known from Friesz et al. (1993) to describe a simultaneous path and departure choice dynamic equilibrium, while (9.129) through (9.134) were shown in Section 9.2.5 to be valid constrained dynamics based on the separable arc delay model. However, with regard to the formal development presented so far in this chapter, formulation (9.135) is conjectural; it remains for us to formally demonstrate that its solutions will in fact be dynamic network user equilibria. This is done in Section 9.6.3.

9.6.1 Nested Operators

To create the desired differential variational inequality, we need first to more fully characterize the delay operators. In light of how the arc exit time functions were defined in Section 9.2.4, it is immediate that path traversal time may be expressed as

$$\begin{aligned} D_p(t) &= \sum_{i=1}^{m(p)} [\tau^p_{a_i}(t) - \tau^p_{a_{i-1}}(t)] \\ &= [\tau^p_{a_1}(t) - t] + [\tau^p_{a_2}(t) - \tau^p_{a_1}(t)] + \ldots + [\tau^p_{a_{m(p)}}(t) - \tau_{a_{m(p)-1}}(t)] \\ &= \tau^p_{a_{m(p)}}(t) - t \quad \forall p \in \mathcal{P} \end{aligned} \tag{9.136}$$

for all $p \in \mathcal{P}$. Moreover, as we have already observed, the exit time functions obey the recursive identities

$$\tau^p_{a_1}(t) = t + D_{a_1}[x_{a_1}(t)] \quad \forall p \in \mathcal{P} \tag{9.137}$$

and

$$\tau^p_{a_i}(t) = \tau^p_{a_{i-1}}(t) + D_{a_i}[x_{a_i}(\tau^p_{a_{i-1}}(t))] \quad \forall p \in \mathcal{P}, i \in [2, m(p)] \tag{9.138}$$

It is easy to see that (9.137) and (9.138) may be used to construct nested delay operators that rapidly become very complex as the number of arcs comprising a path increases, as discussed by Wie et al. (1995). Also, note that, because $D_p(t) = \tau^p_{a_{m(p)}}(t) - t$, we may write the path delay operators in the following form:

$$\begin{aligned} D_p(t, x) &\equiv \text{path delay for departure at time } t \text{ under traffic conditions } x \\ &= \sum_{i=1}^{m(p)} \delta_{a_i p} \Phi_{a_i}(t, x) \end{aligned} \tag{9.139}$$

where the $\Phi_{a_i}(t, x)$ are arc delay operators obeying

$$\begin{aligned} &\Phi_{a_1}(t, x) = D_{a_1}\left[x_{a_1}(t)\right] \\ &\Phi_{a_2}(t, x) = D_{a_2}\left[x_{a_2}\left(t + \Phi_{a_1}\right)\right] \\ &\Phi_{a_3}(t, x) = D_{a_3}\left[x_{a_3}\left(t + \Phi_{a_1} + \Phi_{a_2}\right)\right] \\ &\vdots \\ &\Phi_{a_i}(t, x) = D_{a_i}\left[x_{a_i}\left(t + \Phi_{a_i-1} + \ldots + \Phi_{a_1}\right)\right] = D_{a_i}\left[x_{a_i}\left(t + \sum_{j=1}^{i-1} \Phi_{a_j}\right)\right] \end{aligned} \tag{9.140}$$

We also introduce the asymmetric arrival penalty

$$\Theta\left[t + D_p(t, x) - T_A\right] \tag{9.141}$$

where T_A is the prescribed fixed arrival time and $T_A > t_f$. The arrival penalty operator has the properties

$$t + D_p(t,x) > T_A \Longrightarrow \Theta\left[t + D_p(t,x) - T_A\right] = \chi^L(t,x) > 0 \tag{9.142}$$

$$t + D_p(t,x) < T_A \Longrightarrow \Theta\left[t + D_p(t,x) - T_A\right] = \chi^E(t,x) > 0 \tag{9.143}$$

$$t + D_p(t,x) = T_A \Longrightarrow \Theta\left[t + D_p(t,x) - T_A\right] = 0 \tag{9.144}$$

$$\chi^L(t,x) > \chi^E(t,x) \tag{9.145}$$

for every path $p \in \mathcal{P}$ to reflect the fact that arriving late is more serious than arriving early. Consequently, the effective delay operator for each path $p \in \mathcal{P}$ is

$$F_p(t,h) = D_p(t,x) + \Theta\left[t + D_p(t,x) - T_A\right] \tag{9.146}$$

since the states x (as well as the arc exit flows g) are completely determined by knowledge of h.

9.6.2 The Problem Setting

In Chapter 6 we studied differential variational inequalities that possess state-dependent time shifts, as does the proposed formulation (9.135). In order to apply the formalism developed in Chapter 6, we make the following observations/assumptions:

1. the controls are $g \in \left(L^2\left[t_0, t_f\right]\right)^{n_1}$ and $h \in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{P}|}$ where

$$n_1 = \sum_{p=1}^{|\mathcal{P}|} m(p);$$

2. the state variables are the traffic volumes

$$x_{a_i}^p \qquad \forall p \in \mathcal{P}, i \in [1, m(p)];$$

3. the arc delays

$$D_{a_i}\left(x_{a_i}\right) \qquad \forall p \in \mathcal{P}, i \in [1, m(p)];$$

   appear as explicit time shifts in the flow propagation constraints;
4. the controls g (arc exit flows) are intermediate variables that could be eliminated using the flow propagation constraints, allowing us to look at the states as operators of the form $x(h,t)$, when doing so clarifies our exposition;
5. the state operator $x(h,t)$ is continuous and G-differentiable;
6. each $D_{a_i}(x)$ is continuously differentiable with respect to its own argument;
7. an appropriate Lipschitz condition holds for the effective delay operator $F(t,h)$; and

8. in discussing algorithm convergence, we will assume $F(t,h)$ is strongly monotone with respect to h, although examples of nonmonotonic effective delay operators may be constructed; for other-than-convergence discussions this assumption will not be invoked.

It should be noted that conditions (1) through (6) immediately above present no difficulty for either analysis or computation. Condition (7) is not particularly restrictive. However, condition (8), which is needed to assure convergence of our fixed-point iterative scheme, is known to to be violated in dynamic user equilibrium. Given this remark, it should be evident that the fixed-point algorithm is presently a heuristic algorithm, when applied to the dynamic user equilibrium problem.

9.6.3 Analysis

To facilitate the analysis of the necessary conditions for (9.135), it is helpful to restate that differential variational inequality as the following optimal control problem:

$$\min \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} F_p\left(t,h^*\right) h_p(t)\,dt \tag{9.147}$$

subject to

$$\frac{dx_{a_1}^p(t)}{dt} = h_p(t) - g_{a_1}^p(t) \qquad \left(\lambda_{a_1}^p\right) \ \ \forall p\in\mathcal{P} \tag{9.148}$$

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \left(\lambda_{a_i}^p\right) \ \ \forall\, p\in\mathcal{P}, i\in[2,m(p)] \tag{9.149}$$

$$\frac{dy_{ij}(t)}{dt} = \sum_{p\in\mathcal{P}_{ij}} h_p(t) \qquad \left(\mu_{ij}\right) \ \ \forall (i,j)\in\mathcal{W} \tag{9.150}$$

$$g_{a_1}\left(t + D_{a_1}\left(x_{a_1}\right)\right)\left(1 + D'_{a_1}\left(x_{a_1}\right)\dot{x}_{a_1}\right) - h_p = 0 \qquad \left(\gamma_{a_1}^p\right) \ \ \forall p\in\mathcal{P} \tag{9.151}$$

$$g_{a_i}^p\left(t + D_{a_i}\left(x_{a_i}\right)\right)\left(1 + D'_{a_i}\left(x_{a_i}\right)\dot{x}_{a_i}\right) - g_{a_{i-1}}^p = 0 \ \ \left(\gamma_{a_i}^p\right) \ \ \forall\, p\in\mathcal{P},\ i\in[2,m(p)] \tag{9.152}$$

$$-h_p \le 0 \qquad \left(\rho_{a_0}^p\right) \ \ \forall p\in\mathcal{P} \tag{9.153}$$

$$-g_{a_i}^p(t) \le 0 \qquad \left(\rho_{a_i}^p\right) \ \ \forall p\in\mathcal{P}, i\in[1,m(p)] \tag{9.154}$$

$$-x_{a_i}^p(t) \le 0 \qquad \left(\zeta_{a_i}^p\right) \ \ \forall p\in\mathcal{P}, i\in[1,m(p)] \tag{9.155}$$

$$y_{ij}\left(t_f\right) = Q_{ij} \qquad \left(\nu_{ij}\right) \ \ \forall (i,j)\in\mathcal{W} \tag{9.156}$$

$$y_{ij}(t_0) = 0 \qquad \left(\nu_{ij}\right) \ \ \forall (i,j)\in\mathcal{W} \tag{9.157}$$

$$x(t_0) = x_0 \qquad y(t_0) = y_0 \tag{9.158}$$

We call this optimal control problem $OCP(F^*, \Gamma)$. Some remarks are in order concerning this formulation:

1. Any solution of (9.135) must be a solution of (9.147) through (9.158). This is because (9.135) is a necessary condition for $OCP(F^*, \Gamma)$;
2. $OCP(F^*, \Gamma)$ cannot be used for computation as stated because its articulation assumes knowledge of the dynamic user equilibrium departure rates h^* that generate the state vector x^*. Rather, $OCP(F^*, \Gamma)$ is a mathematical convenience for analyzing the necessary conditions of $DVI\ (F, \Gamma)$, allowing us to use the minimum principal and other necessary conditions of optimal control theory;
3. $OCP(F^*, \Gamma)$ is an optimal control problem with state-dependent time shifts;
4. the variables in parentheses next to the dynamics (9.148), (9.149), and (9.150) are the traditional adjoint variables of optimal control theory, and the variables in parentheses next to the remaining constraints are dual variables;
5. in (9.150) we have used the standard device for treating isoperimetric constraints of introducing a new state variable y_{ij} and terminal time condition $y_{ij}\left(t_f\right) - y_{ij}\left(t_0\right) = Q_{ij}$ to replace each flow conservation constraint

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p\left(t\right) dt = Q_{ij}\ ;$$

6. problem $OCP(F^*, \Gamma)$ has linear dynamics and a linear objective, so its Hamiltonian is linear in controls (and states); and
7. the minimum principle is a convex mathematical program, the minimization of a linear objective subject to linear constraints, that may be solved by inspection.

With the above features in mind, let us turn to the question of whether the suggested differential variational inequality does in fact describe a dynamic user equilibrium.

In particular, let us next assume we have a solution of $DVI\ (F, \Lambda_1)$. We wish to show that this solution is a dynamic network user equilibrium. Because of the observations made previously, it is enough to show that the necessary conditions for (9.147) through (9.158) are a description of dynamic network user equilibrium. Our analysis will be facilitated by the shorthand

$$\tilde{g}_{a_i}^p \equiv g_{a_i}^p\left(t + D_{a_i}\left[x_{a_i}\right]\right) \qquad \forall\ p \in \mathcal{P}, i \in [1, m\left(p\right)] \tag{9.159}$$

This allows us to write the augmented Hamiltonian H_1 for optimal control problem (9.147) through (9.158) as

$$\begin{aligned} H_1 = & \sum_{p \in \mathcal{P}} F_p\left(t, h^*\right) h_p + \sum_{p \in \mathcal{P}} \lambda_{a_1}^p \left(h_p - g_{a_1}^p\right) + \sum_{p \in \mathcal{P}} \sum_{i \in [2, m(p)]} \lambda_{a_i}^p \left(g_{a_{i-1}}^p - g_{a_i}^p\right) \\ & + \sum_{(k,l) \in \mathcal{W}} \mu_{kl} \sum_{p \in \mathcal{P}_{kl}} h_p + \sum_{p \in \mathcal{P}} \gamma_{a_1}^p \left[\tilde{g}_{a_i}^p \cdot \left(1 + D'_{a_1} \dot{x}_{a_1}\right) - h_p\right] \\ & + \sum_{p \in \mathcal{P}} \sum_{i \in [2, m(p)]} \gamma_{a_i}^p \left[\tilde{g}_{a_i}^p \cdot \left(1 + D'_{a_i} \dot{x}_{a_i}\right) - g_{a_{i-1}}^p\right] \end{aligned}$$

$$+\sum_{(k,l)\in\mathcal{W}} v_{kl}\left(t_f\right)\left(Q_{kl}-y_{kl}\left(t_f\right)\right)$$
$$-\sum_{p\in\mathcal{P}}\rho^p_{a_0}h_p-\sum_{p\in\mathcal{P}}\sum_{i\in[1,m(p)]}\rho^p_{a_i}g^p_{a_i}+\sum_{p\in\mathcal{P}}\sum_{i\in[1,m(p)]}\zeta^p_{a_i}x^p_{a_i} \tag{9.160}$$

In taking partial derivatives of H_1, we will treat the $\tilde{g}^p_{a_i}$ as though they were separate controls, temporarily ignoring their relationship to the $g^p_{a_i}$ and the x_{a_i}. This is possible because we employ a specific form of the minimum principle, taken from Chapter 4, for optimal control problems with time shifts; the minimum principle accounts for the fact that the $\tilde{g}^p_{a_i}$ are future values of the $g^p_{a_i}$ determined by the state variables $x^p_{a_i}$. In other words, the formulae we will use for differentiating the Hamiltonian will correct for treating the $\tilde{g}^p_{a_i}$ as independent variables when analyzing the minimum principle.

We are going to need the following partial derivatives of H_1:

$$\frac{\partial H_1}{\partial h_p}=F_p\left(t,h^*\right)+\mu_{ij}+\lambda^p_{a_1}-\gamma^p_{a_1}-\rho^p_{a_0}\qquad\forall\ (i,j)\in\mathcal{W},p\in\mathcal{P}_{ij} \tag{9.161}$$

$$\frac{\partial H_1}{\partial g^p_{a_i}}=-\lambda^p_{a_i}+\lambda^p_{a_{i+1}}-\rho^p_{a_i}-\gamma^p_{a_{i+1}}\qquad\forall\ p\in\mathcal{P},i\in[1,m\left(p\right)-1] \tag{9.162}$$

$$\frac{\partial H_1}{\partial g^p_{a_m}}=-\lambda^p_{a_{m(p)}}-\rho^p_{a_{m(p)}}\qquad\forall\ p\in\mathcal{P} \tag{9.163}$$

$$\frac{\partial H_1}{\partial \tilde{g}^p_{a_i}}=\gamma^p_{a_i}\left(1+D'_{a_i}\dot{x}_{a_i}\right)\qquad\forall\ p\in\mathcal{P},i\in[1,m\left(p\right)] \tag{9.164}$$

$$\frac{\partial H_1}{\partial x^p_{a_i}}=\zeta^p_{a_i}\qquad\forall\ p\in\mathcal{P},i\in[1,m\left(p\right)-1] \tag{9.165}$$

We are now ready to apply the actual necessary conditions for optimal control with state-dependent time shifts derived in the Chapter 4.

Because we have priced-out all constraints, the following form of the minimum principle applies:

$$\left.\frac{\partial H_1}{\partial h_p}\right|_{t^p_0}=0\qquad\forall\ (i,j)\in\mathcal{W},p\in\mathcal{P}_{ij} \tag{9.166}$$

$$\left.\frac{\partial H_1}{\partial g^p_{a_i}}\right|_{t^p_i}+\left[\frac{\partial H_1}{\partial \tilde{g}^p_{a_i}}\cdot\frac{1}{\left(1+D'_{a_i}\dot{x}_{a_i}\right)}\right]_{t^p_{i-1}}=0\qquad\forall\ p\in\mathcal{P},\ i\in[1,m\left(p\right)] \tag{9.167}$$

where $[.]_{t^p_{i-1}}$ means that the argument of this operator is to be evaluated at t^p_{i-1}, the moment in time of entering arc $a_i\in p$ when

$$t^p_i=t^p_{i-1}+D_{a_i}\left[x_{a_i}\left(t^p_{i-1}\right)\right]$$

is the moment in time of exiting arc $a_i \in p$. Taking t_0^p as the time of departure from the origin, it is immediate from (9.161) and (9.166) that

$$F_p\left(t_0^p, x^*\right) + \mu_{ij} + \lambda_{a_1}^p\left(t_0^p\right) - \gamma_{a_1}^p\left(t_0^p\right) - \rho_{a_0}^p\left(t_0^p\right) = 0 \qquad \forall\ (i,j) \in \mathcal{W}, p \in \mathcal{P}_{ij} \tag{9.168}$$

For condition (9.167), we consider the cases $i \in [1, m(p) - 1]$ and $i = m(p)$ separately. Using the derivatives (9.162) and (9.163) in (9.167), gives

$$\lambda_{a_{i+1}}^p\left(t_i^p\right) - \lambda_{a_i}^p\left(t_i^p\right) - \gamma_{a_{i+1}}^p\left(t_i^p\right) - \rho_{a_i}^p\left(t_i^p\right) + \tilde{\gamma}_{a_i}^p\left(t_i^p\right) = 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p) - 1] \tag{9.169}$$

$$-\lambda_{a_{m(p)}}^p\left(t_m^p\right) - \rho_{a_{m(p)}}^p\left(t_m^p\right) + \tilde{\gamma}_{a_{m(p)}}^p\left(t_m^p\right) = 0 \qquad \forall\ p \in \mathcal{P} \tag{9.170}$$

In these last two expressions we have employed the notation

$$\tilde{\gamma}_{a_i}^p\left(t_i^p\right) \equiv \gamma_{a_i}^p\left(t_{i-1}^p\right) \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] \tag{9.171}$$

where t_i^p and t_{i-1}^p are as defined previously. Also among the necessary conditions are the following complementary slackness conditions:

$$\rho_{a_0}^p h_p = 0 \qquad \rho_{a_0}^p \geq 0 \qquad \forall\ p \in \mathcal{P} \tag{9.172}$$

$$\rho_{a_i}^p g_{a_i}^p = 0 \qquad \rho_{a_i}^p \geq 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] \tag{9.173}$$

$$\zeta_{a_i}^p x_{a_i}^p = 0 \qquad \zeta_{a_i}^p \geq 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] \tag{9.174}$$

where it is understood that these conditions apply for all $t \in \left[t_0, t_f\right]$.

We further observe that the necessary conditions for the time evolution of the $\lambda_{a_i}^p$, known as adjoint equations, are

$$\begin{aligned} -\frac{d\lambda_{a_i}^p}{dt} &= \frac{\partial H_1}{\partial x_{a_i}^p} - \frac{d}{dt}\frac{\partial H_1}{\partial \dot{x}_{a_i}^p} \\ &= \zeta_{a_i}^p + \left[\frac{\partial}{\partial x_{a_i}^p}\sum_q\sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \cdot \left(1 + D'_{a_j}\dot{x}_{a_j}\right)\right] \\ &\quad - \frac{d}{dt}\frac{\partial}{\partial \dot{x}_{a_i}^p}\sum_q\sum_j \left(\gamma_{a_j}^q \tilde{g}_{a_j}^q D'_{a_j}\dot{x}_{a_j}\right) \\ &= \zeta_{a_i}^p + \sum_q\sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_i}^p}\dot{x}_{a_j} - \frac{d}{dt}\sum_q\sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q D'_{a_j}\frac{\partial \dot{x}_{a_j}}{\partial \dot{x}_{a_i}^p} \end{aligned} \tag{9.175}$$

Noting that

$$\frac{\partial \dot{x}_{a_j}}{\partial \dot{x}_{a_i}^p} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases}$$

we have

$$\begin{aligned} -\frac{d\lambda_{a_i}^p}{dt} &= \zeta_{a_i}^p + \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_i}^p} \dot{x}_{a_j} - \frac{d}{dt} \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q D'_{a_j} \\ &= \zeta_{a_i}^p + \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_i}^p} \dot{x}_{a_j} - \frac{d}{dt} \sum_p \sum_k \frac{dx_{a_k}^p}{dt} \frac{\partial}{\partial x_{a_k}^p} \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q D'_{a_j} \end{aligned}$$

which results in

$$\begin{aligned} -\frac{d\lambda_{a_i}^p}{dt} &= \zeta_{a_i}^p + \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_i}^p} \dot{x}_{a_j} - \sum_r \frac{dx_{a_j}^r}{dt} \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_j}^p} \\ &= \zeta_{a_i}^p + \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_i}^p} \dot{x}_{a_j} - \sum_q \sum_j \gamma_{a_j}^q \tilde{g}_{a_j}^q \frac{\partial D'_{a_j}}{\partial x_{a_j}^p} \dot{x}_{a_j} \\ &= \zeta_{a_i}^p \qquad \forall p \in \mathcal{P}, i \in [1, m(p)] \end{aligned} \tag{9.176}$$

when $k = j$.

We now define an *open path* $p = \{a_1, a_2, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_{m(p)}\} \in \mathcal{P}_{kl}$ by the conditions

$$\left.\begin{aligned} & g_{a_0}^p\left(t_0^p\right) \equiv h_p\left(t_0^p\right) > 0 \\ & g_{a_i}^p\left(t_{i-1}^p\right) > 0 \qquad \forall\, i \in [1, m(p)] \\ & g_{a_i}^p\left(t_i^p\right) > 0 \qquad \forall\, i \in [1, m(p)] \\ & x_{a_i}^p(t) > 0 \qquad \forall\, t \in \left[t_{i-1}^p, t_i^p\right], i \in [1, m(p)] \end{aligned}\right\} \tag{9.177}$$

Note that the complementary slackness conditions (9.172), (9.173), and (9.174) require that for an open path

$$\left.\begin{aligned} & \rho_{a_0}^p\left(t_0^p\right) = 0 \\ & \rho_{a_i}^p\left(t_{i-1}^p\right) = \rho_{a_i}^p\left(t_i^p\right) = 0 \qquad \forall\, i \in [1, m(p)] \\ & \zeta_{a_i}^p(t) = 0 \qquad \forall\, t \in \left[t_{i-1}^p, t_i^p\right], i \in [1, m(p)] \end{aligned}\right\} \tag{9.178}$$

The identities (9.169), (9.170), and (9.176) become

$$\lambda_{a_{i+1}}^p\left(t_i^p\right) - \gamma_{a_{i+1}}^p\left(t_i^p\right) = \lambda_{a_i}^p\left(t_i^p\right) - \tilde{\gamma}_{a_i}^p\left(t_i^p\right) \quad \forall\, i \in [1, m(p) - 1],\ t \in \left[t_{i-1}^p, t_i^p\right] \tag{9.179}$$

$$\lambda_{a_m}^p\left(t_m^p\right) - \tilde{\gamma}_{a_m}^p\left(t_m^p\right) = 0 \tag{9.180}$$

$$\frac{d\lambda_{a_i}^p(t)}{dt} = 0 \qquad \forall\, t \in \left[t_{i-1}^p, t_i^p\right],\ i \in [1, m(p)] \tag{9.181}$$

for an open path described by (9.177) and (9.178)

We see from (9.181) that on the last arc of the open path which we are considering

$$\frac{d\lambda_{a_m}^p(t)}{dt} = 0 \qquad t \in \left[t_{m-1}^p, t_m^p\right] \tag{9.182}$$

That is, $\lambda_{a_m}^p$ is time invariant during the time flow traverses arc a_m, so that

$$\lambda_{a_m}^p\left(t_{m-1}^p\right) = \lambda_{a_m}^p\left(t_m^p\right) \tag{9.183}$$

It then follows from (9.180) and (9.183) that

$$\lambda_{a_m}^p\left(t_{m-1}^p\right) = \lambda_{a_m}^p\left(t_m^p\right) = \tilde{\gamma}_{a_m}^p\left(t_m^p\right) = \gamma_{a_m}^p\left(t_{m-1}^p\right) \tag{9.184}$$

where we have used the identity $\tilde{\gamma}_{a_m}^p\left(t_m^p\right) = \gamma_{a_m}^p\left(t_{m-1}^p\right)$ obtained from (9.171). It is immediate from (9.184) that

$$\lambda_{a_m}^p\left(t_{m-1}^p\right) - \gamma_{a_m}^p\left(t_{m-1}^p\right) = 0 \tag{9.185}$$

Using the recursive formula (9.179) together with (9.185), we have

$$\lambda_{a_m}^p\left(t_{m-1}^p\right) - \gamma_{a_m}^p\left(t_{m-1}^p\right) = \lambda_{a_{m-1}}^p\left(t_{m-1}^p\right) - \tilde{\gamma}_{a_m}^p\left(t_{m-1}^p\right) = 0 \tag{9.186}$$

Additionally, we know from (9.181) that

$$\frac{d\lambda_{a_{m-1}}^p(t)}{dt} = 0 \qquad t \in \left[t_{m-2}^p, t_{m-1}^p\right] \tag{9.187}$$

which of course means that $\lambda_{a_{m-1}}^p(t)$ is time invariant for $t \in \left[t_{m-2}^p, t_{m-1}^p\right]$, so that

$$\lambda_{a_{m-1}}^p\left(t_{m-2}^p\right) = \lambda_{a_{m-1}}^p\left(t_{m-1}^p\right) \tag{9.188}$$

This last fact together with the identity $\tilde{\gamma}_{a_m}^p\left(t_{m-1}^p\right) = \gamma_{a_{m-1}}^p\left(t_{m-2}^p\right)$ obtained from (9.171) means that (9.186) implies

$$\lambda_{a_{m-1}}^p\left(t_{m-2}^p\right) - \gamma_{a_{m-1}}^p\left(t_{m-2}^p\right) = 0 \tag{9.189}$$

Proceeding inductively in this fashion, we arrive at the result

$$\lambda_{a_1}^p\left(t_0^p\right) - \gamma_{a_1}^p\left(t_0^p\right) = 0 \tag{9.190}$$

for any open path.

To conclude our analysis, we must exploit the transversality conditions

$$\frac{\partial \sum_{(i,j)\in\mathcal{W}} v_{ij}\left(t_f\right)\left(Q_{ij}-y_{ij}\left(t_f\right)\right)}{\partial y_{kl}\left(t_f\right)}=\mu_{kl}\Longrightarrow \mu_{kl}=-v_{kl}\left(t_f\right) \qquad \forall\ (k,l)\in\mathcal{W} \tag{9.191}$$

Conditions (9.191) together with the implications of complementary slackness for an open path (9.178) and the identity (9.190) allow us to extract from (9.168) the following property of an open path p:

$$\begin{aligned} h_p^*\left(t_0^p\right)>0,\ p\in\mathcal{P}_{kl} &\Longrightarrow \lambda_{a_1}^p\left(t_0^p\right)-\gamma_{a_1}^p\left(t_0^p\right)-\rho_{a_0}^p\left(t_0^p\right)=0 \\ &\Longrightarrow F_p\left(t_0^p,h^*\right)=v_{kl}\left(t_f\right) \end{aligned} \tag{9.192}$$

which is immediately recognized as the fundamental condition for a dynamic network user equilibrium. Moreover, we also see that

$$F_p\left(t_0^p,h^*\right)>v_{kl}\left(t_f\right),\ p\in\mathcal{P}_{kl}\Longrightarrow h_p^*\left(t_0^p\right)=0 \tag{9.193}$$

since the only alternative implication of $F_p\left(t_0^p,h^*\right)>v_{kl}\left(t_f\right)$ would be $h_p^*\left(t_0^p\right)>0$ which would directly contradict (9.192); thus, we have established that any solution of the differential variational inequality formulation we have proposed is a dynamic user equilibrum. That is, the preceeding analysis has provided, as we intended, a constructive proof of the following theorem:

Theorem 9.4. *Any solution of $DVI(F,\Gamma)$ given by (9.135) is a dynamic network user equilibrium relative to departure time and path choice.*

9.6.4 Computation with Endogenous Delay Operators

We know from Chapter 6 that one approach to the numerical solution of differential variational inequalities is to convert them to fixed-point problems involving linear-quadratic optimal control problems as subproblems; it is significant that such an approach does not require the complicated effective path delay operators encountered in DUE to be differentiated. We comment, however, that the convergence of the fixed-point algorithm described in this chapter when applied to the dynamic user equilibrium problem may not be assured when using results from Chapter 6 because the effective delay operator will not generally satisfy the monotonicity condition invoked to establish convergence.

9.6.4.1 Dealing with Time Shifts

A critical hurdle to clear in solving the differential variational inequality representation of DUE is that of dealing with the state-dependent time shifts

intrinsic to the flow propagation constraints, when the delay operators are endogenously determined. We deal with this challenge by employing an implicit fixed-point computational scheme that, for each main iteration, approximates any decision variable with shifted argument as a pure, continuous function of time. It is easiest to understand this scheme in the abstract without the complicated notation that surrounds the differential inequality representation of dynamic user equilibrium. To do this, let us suppose one is faced at iteration ℓ with the need to evaluate the following abstract control with shifted argument

$$u\left[t + \Delta\left(x\left(t\right)\right)\right] \tag{9.194}$$

for which t denotes continuous time, x denotes an abstract state variable, and the time shift Δ is state-dependent. Further suppose that in iteration $\ell - 1$ we found the continuous time results

$$\Delta\left(x\left(t\right)\right) \approx \sum_{j=0}^{r} \alpha_j^{\ell-1} \cdot (t)^j \tag{9.195}$$

$$u\left(t\right) \approx \sum_{k=0}^{s} \beta_k^{\ell-1} \cdot (t)^k \tag{9.196}$$

where s, r, $\alpha_j^{\ell-1}$, and $\beta_k^{\ell-1}$ are arbitrary names of parameters and the polynomial forms are meant merely to be illustrative. For such a circumstance we form the following approximation of the shifted control

$$u^{\ell}\left(t + \Delta\right) \approx \sum_{k=0}^{s} \beta_k^{\ell-1} \cdot \left(t + \sum_{j=0}^{r} \alpha_j^{\ell-1} \left(t\right)^j \right)^k \equiv \tilde{u}^{\ell}\left(t\right) \tag{9.197}$$

Clearly, (9.197) is a pure function of time constructed from information acquired in iteration $\ell - 1$ and meant for use in iteration ℓ. It is important to note that the continuous-time forms (9.195) and (9.196) may be either the direct result of a continuous-time analysis or they may be polynomials fit to the output of a discrete time approximation. In either case it is possible to form the approximation $\tilde{u}^{\ell}$. As the main algorithm proceeds through iterations $\ell + 1$, $\ell + 2$, and so forth, one is refining the estimate of $u\left(t + \Delta\right)$. If the algorithm's action on the time shifted variable is denoted by the abstract operator Y, then the proposed approach may be expressed as that of seeking a repesentation $\tilde{u}^*$ that is a fixed point according to

$$\tilde{u}^* = Y\left(\tilde{u}^*\right)$$

where

$$\lim_{\ell \longrightarrow \infty} \tilde{u}^{\ell} = \tilde{u}^*$$

although the operator Y will not generally be expressible in closed form.

9.6.4.2 *A Numerical Example*

As a numerical example of application of the fixed-point algorithm for differential variational inequalities presented in Chapter 6, augmented by the implicit fixed-point scheme of Section 9.6.4.1 immediately above, let us now consider the five-arc, four-node network shown in Figure 9.1 and considered by Friesz and Mookherjee (2006). The forward-star array is

Arc Name	From Node	To Node
a_1	1	2
a_2	1	3
a_3	2	3
a_4	2	4
a_5	3	4

The arc delay functions considered are

$$D_{a_1} = \frac{1}{2} + \frac{1}{70}x_{a_1} \quad D_{a_4} = 1 + \frac{1}{150}x_{a_4}$$

$$D_{a_2} = 1 + \frac{1}{150}x_{a_2} \quad D_{a_5} = \frac{1}{2} + \frac{1}{100}x_{a_5}$$

$$D_{a_3} = \frac{1}{2} + \frac{1}{100}x_{a_3}$$

There is a travel demand of $Q = 75$ units from node 1 (the single origin) to node 4 (the single destination). There are three paths connecting origin-destination pair $(1, 4)$; specifically, we have

$$\mathcal{P}_{14} = \{p_1, p_2, p_3\}$$
$$p_1 = \{a_1, a_4\}$$

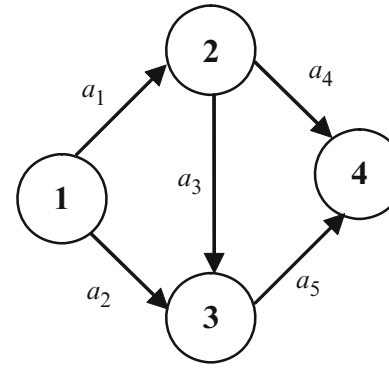


Fig. 9.1 Five-arc, four-node traffic network

$$p_2 = \{a_2, a_5\}$$
$$p_3 = \{a_1, a_3, a_5\}$$

The controls (path flows and arc exit flows) and states (path-specific arc volumes) associated with this network are

Paths	path flows	arc exit flows	arc volumes
p_1	h_{p_1}	$g_{a_1}^{p_1}, g_{a_4}^{p_1}$	$x_{a_1}^{p_1}, x_{a_4}^{p_1}$
p_2	h_{p_2}	$g_{a_2}^{p_2}, g_{a_5}^{p_2}$	$x_{a_2}^{p_2}, x_{a_5}^{p_2}$
p_3	h_{p_2}	$g_{a_1}^{p_3}, g_{a_3}^{p_3}, g_{a_5}^{p_3}$	$x_{a_1}^{p_3}, x_{a_3}^{p_3}, x_{a_5}^{p_3}$

The path flows and arc exit flows are bounded from above by 15 and from below by zero:

$$0 \leq h_p \leq 15 \quad \forall p \in \mathcal{P}$$
$$0 \leq g_{a_i}^p \leq 15 \quad \forall p \in \mathcal{P}, i \in [1, m(p)]$$

The initial time is t_0, the terminal time is $t_f = 10$, and the time interval considered is $[0, 10]$, while the desired arrival time is $t_A = 5$. We employ the symmetric early/late arrival penalty

$$F\left[t + D_p(x,t) - t_A\right] = \left[t + D_p(x,t) - t_A\right]^2$$

Furthermore, without any loss of generality, we take

$$x_{a_i}^p(t_0) = 0 \qquad \forall p \in \mathcal{P}, i \in [1, m(p)]$$

The fixed-point stopping tolerance will be set at

$$\varepsilon = 0.01$$

To assist convergence we choose the free parameter of the fixed-point formulation to obey

$$\alpha = \frac{1}{k}$$

where k is the fixed-point iteration counter.

We forgo the detailed symbolic statement of this example and, instead, provide numerical results in graphical form for the solution after 77 iterations of the fixed-point algorithm. Figures 9.2, 9.3, and 9.4 depict path flows and arc exit flows for paths p_1, p_2, and p_3 defined above. Cumulative traffic volumes on the network's 5 different arcs are plotted against time in Figure 9.5 where

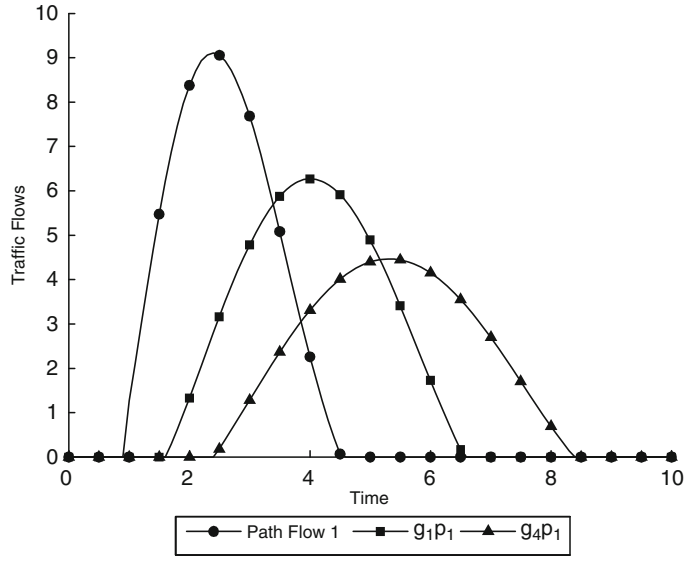


Fig. 9.2 Path and arc exit flows for path p_1

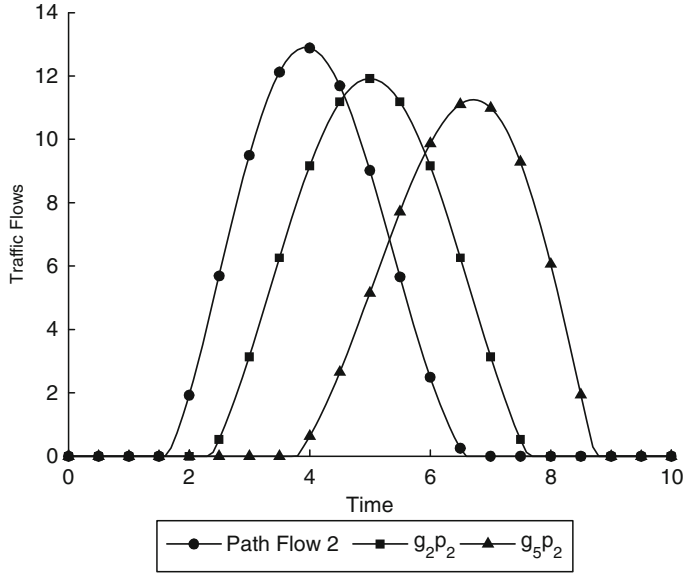


Fig. 9.3 Path and arc exit flows for path p_2

$$x_{a_1}(t) = x_{a_1}^{p_1}(t) + x_{a_1}^{p_3}(t)$$
$$x_{a_2}(t) = x_{a_2}^{p_2}(t)$$
$$x_{a_3}(t) = x_{a_3}^{p_3}(t)$$

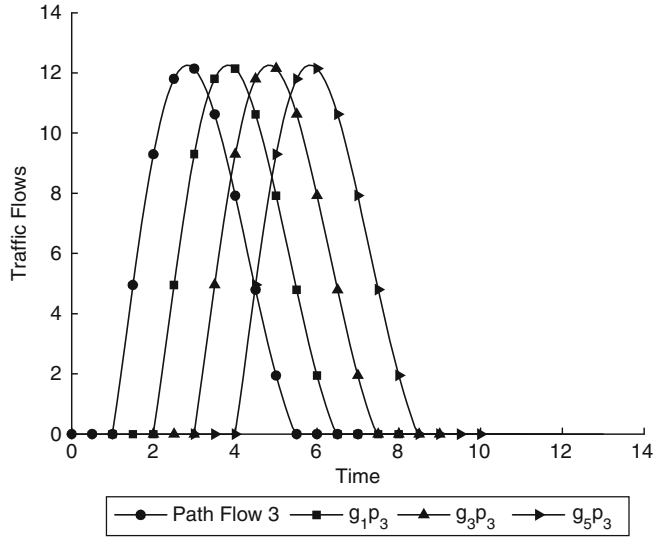


Fig. 9.4 Path and arc exit flows for path p_3

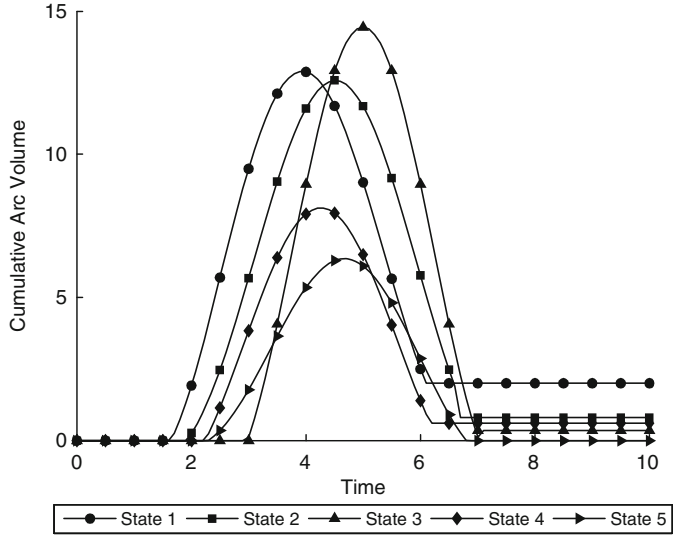


Fig. 9.5 Cumulative traffic volume on each arc

$$x_{a_4}(t) = x_{a_4}^{p_1}(t)$$
$$x_{a_5}(t) = x_{a_5}^{p_2}(t) + x_{a_5}^{p_3}(t)$$

for all $t \in [0, 10]$. Figure 9.5 presents arc volumes derived from activity on all paths and corresponding to the departure rates. When we compare the effective path

delay operator (9.146) with path flow (departure rate) by plotting both for the same time scale, Figures 9.6 and 9.7 are obtained. These figures show that departure rate peaks when the associated effective path delay achieves a local minimum, thereby demonstrating that an user equilibrium has been found.

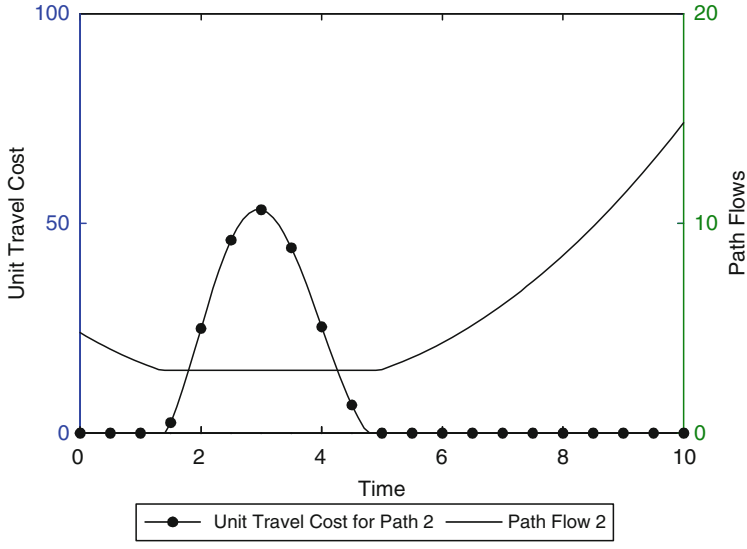


Fig. 9.6 Comparison of path flow and unit travel cost for path p_2

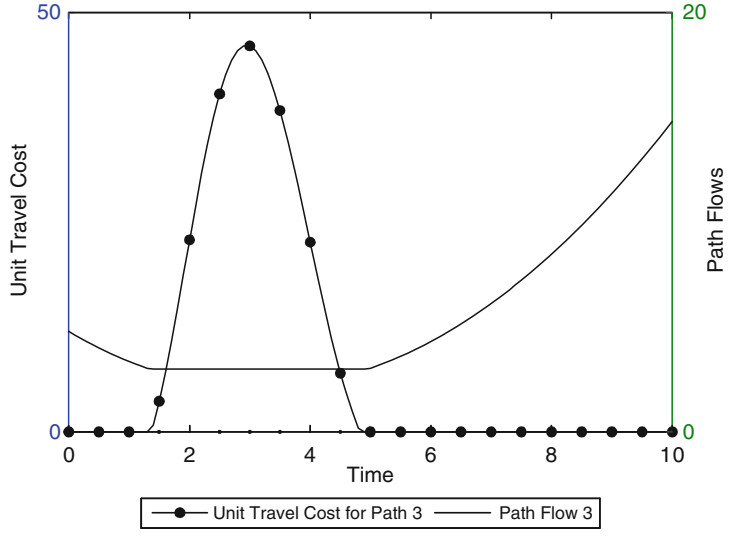


Fig. 9.7 Comparison of path flow and unit travel cost for path p_3

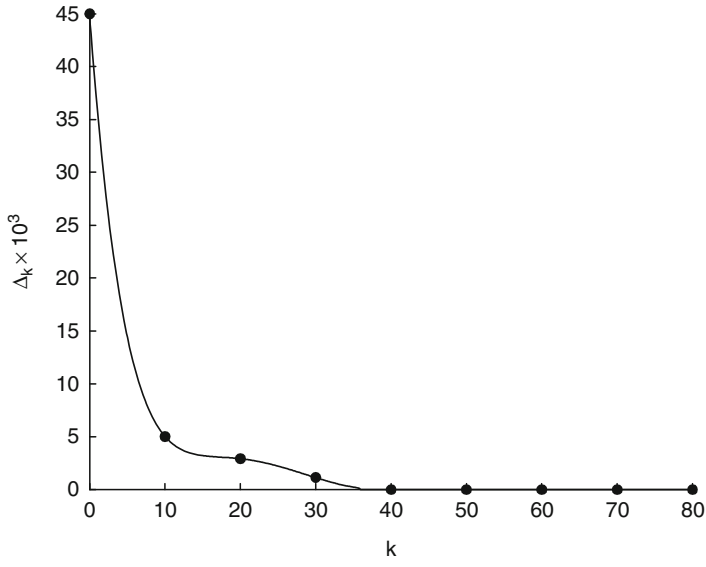


Fig. 9.8 Control fluctuations from one iteration to the next: (Δ_k) vs. the iteration counter (k)

In Figure 9.8 the relative change in exit flows from one iteration to the next, expressed as

$$\Delta_k = \left\| g^k - g^{k-1} \right\| \leq \varepsilon,$$

is plotted against the iteration counter k. It is worth noting that for this particular example even though $\Delta_1 \approx 4 \times 10^4$, the next several iterations see a very rapid decrease of Δ_k. The run time for this example is less than 3 minutes using a generic desktop computer with dual Intel Xeon processors and 2 GB of RAM. The computer code for the fixed-point algorithm is written in MatLab 6.5 and calls a gradient projection subroutine for which control, state and adjoint variables are determined in continuous time, as explained in our discussion of the projected gradient algorithm in Chapter 4. In other words, the solution of this example is achieved using continuous-time methods for both the main fixed-point problem and the projection subproblems.

9.7 Conclusions

In this chapter, we have seen two formulations of the dynamic user equilibrium problem. One is based on exogenous effective delay operators, and the other endogenizes those operators. Each of these models may be expressed as a differential

variational inequality. We have seen how the differential variational inequality formulations may be solved using the methods developed in previous chapters. We also wish to comment in closing that there are several extensions of the models we have reported in this chapter that can be made more or less immediately. These involve the introduction of:

1. nonseparable arc delay functions to model link and nodal interactions associated with turning phenomena and the presence of multiple modes and classes of travelers;
2. elastic travel demand submodels whereby travel demand varies with effective delay;
3. departure and arrival time windows; and
4. stochastic considerations wherein model parameters or unobserved components of generalized travel impedance are governed by probability distributions.

Although these extensions are straightforward from a mathematical point of view, their presentation is quite tedious and is not included here.

9.8 Exercises

1. What properties of dynamic traffic systems make it necessary to consider measure-theoretic issues in defining and constructing mathematical representations of dynamic user equilibrium?
2. Given an argument for the choice of $L^2\left[t_0, t_f\right]$ and $\mathcal{H}^1\left[t_0, t_f\right]$ as fundamental spaces for the study of dynamic traffic assignment.
3. Introduce regularity conditions that allow the existence of a dynamic user equilibrium with fixed travel demand to be proven. Give that proof.
4. Extend the definition of dynamic user equilibrium and its representation as a differential variational inequality to the case of elastic and time-dependent travel demand:
$$Q_{ij}(v, t)$$
where
$$v = \left(v_{ij} : (i, j) \in \mathcal{W}\right)$$
is the essential infimum of effective path delay $F(t, h)$.
5. Attempt a proof of existence for the model you have developed in response to Question 4 above.
6. Pick two methods of solving differential variational inequalities, other than the fixed-point algorithm. Compare and contrast these algorithms to each other and to the fixed-point algorithm when used to solve the dynamic user equilibrium problem with fixed demand.

7. What are the obstacles to proving convergence of a fixed-point algorithm for the dynamic user equilibrium problem in continuous time? Elaborate on and conjecture about how these might be overcome.
8. If the version of the fixed-point algorithm presented in this chapter converges, may we be certain it has computed a *bona fide* dynamic user equilibrium?
9. Numerically solve the following dynamic user equilibium problem using the fixed-point algorithm. The arc delay functions are

$$D_{a_i} = A_i + B_i x_{a_i}$$

The network is described by the array

From Node	To Node	Arc Name	A_i	B_i
1	2	a_3	1.0	5.0
1	4	a_1	1.0	5.0
2	3	a_6	1.0	5.0
2	5	a_4	1.0	5.0
4	2	a_2	1.0	5.0
5	3	a_5	1.0	5.0

There are two origin-destination pairs:

$$\mathcal{W} = \{(1,3),(2,3)\}$$

There are six paths for the two origin-destination pairs, namely:

$$\begin{aligned} p_1 &= \{3,6\} \\ p_2 &= \{1,2,6\} \\ p_3 &= \{1,2,4,5\} \\ p_4 &= \{3,4,5\} \\ p_5 &= \{6\} \\ p_6 &= \{4,5\} \end{aligned}$$

Each origin-destination pair is assumed to have a fixed travel demand:

$$\begin{aligned} Q_{13} &= 100 \\ Q_{23} &= 100 \end{aligned}$$

The desired arrival time of T_A = 9:00 AM for all travelers and both origin-destination pairs. The period of analysis is 7:00 AM to 10:00 AM. You may assume any other data needed for your calculations.

List of References Cited and Additional Reading

Adamo, V., V. Astarita, M. Florian, M. Mahut, and J. H. Wu (1998). A framework for introducing spillback in link-based dynamic network loading models. presented at tristan iii, san juan, puerto rico, june.

Astarita, V. (1995). Flow propagation description in dynamic network loading models. Y.J Stephanedes, F. Filippi, (Eds). *Proceedings of IV international conference of Applications of Advanced Technology in Transportation(AATT)*, 599–603.

Bernstein, D., T. L. Friesz, R. L. Tobin, and B. W. Wie (1993). A variational control formulation of the simultaneous route and departure-time equilibrium problem. *Proceedings of the International Symposium on Transportation and Traffic Theory*, 107–126.

Bliemer, M. and P. Bovy (2003). Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem. *Transportation Research Part B 37*(6), 501–519.

Carey, M. (1986). A constraint qualification for a dynamic traffic assignment problem. *Transportation Science 20*(1), 55–58.

Carey, M. (1987). Optimal time-varying flows on congested networks. *Operations Research 35*(1), 58–69.

Carey, M. (1992). Nonconvexity of the dynamic traffic assignment problem. *Transportation Research 26B*(2), 127–132.

Carey, M. (1995). Dynamic congestion pricing and the price of fifo. In N. H. Gartner and G. Improta (Eds.), *Urban Traffic Networks*, pp. 335–350. New York: Springer-Verlag.

Daganzo, C. (1994). The cell transmission model. Part I: A simple dynamic representation of highway traffic. *Transportation Research B 28*(4), 269–287.

Daganzo, C. (1995). The cell transmission model. Part II: Network traffic. *Transportation Research Part B 29*(2), 79–93.

Friesz, T., D. Bernstein, and R. Stough (1996). Dynamic systems, variational inequalities, and control-theoretic models for predicting time-varying urban network flows. *Transportation Science 30*(1), 14–31.

Friesz, T., D. Bernstein, Z. Suo, and R. Tobin (2001). Dynamic network user equilibrium with state-dependent time lags. *Networks and Spatial Economics 1(3/4)*, 319–347.

Friesz, T., C. Kwon, and D. Bernstein (2007). Analytical dynamic traffic assignment models. In D. A. Hensher and K. J. Button (Eds.), *Handbook of Transport Modelling* (2nd ed.). New York: Pergamon.

Friesz, T., J. Luque, R. Tobin, and B. Wie (1989). Dynamic network traffic assignment considered as a continuous-time optimal control problem. *Operations Research 37*(6), 893–901.

Friesz, T. and R. Mookherjee (2006). Solving the dynamic network user equilibrium problem with state-dependent time shifts. *Transportation Research Part B 40*, 207–229.

Friesz, T., R. Tobin, D. Bernstein, and Z. Suo (1995). Proper flow propagation constraints which obviate exit functions in dynamic traffic assignment. *INFORMS Spring National Meeting, Los Angeles, April 23 26*.

Friesz, T. L., D. Bernstein, T. Smith, R. Tobin, and B. Wie (1993). A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research 41*, 80–91.

Friesz, T. L., P. A. Viton, and R. L. Tobin (1985). Economic and computational aspects of freight network equilibrium: a synthesis. *Journal of Regional Science 25*, 29–49.

Halmos, P. (1974). *Measure Theory*. New York: Springer-Verlag.

Li, Y., S. Waller, and T. Ziliaskopoulos (2003). A decomposition scheme for system optimal dynamic traffic assignment models. *Networks and Spatial Economics 3*(4), 441–455.

Lo, H. and W. Szeto (2002). A cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transportation Research Part B 36*(5), 421–443.

Merchant, D. and G. Nemhauser (1978a). A model and an algorithm for the dynamic traffic assignment problems. *Transportation Science 12*(3), 183–199.

Merchant, D. and G. Nemhauser (1978b). Optimality conditions for a dynamic traffic assignment model. *Transportation Science 12*(3), 200–207.

Nie, Y. and H. M. Zhang (2010). Solving the dynamic user optimal assignment problem considering queue spillback. *Networks and Spatial Economics 10*(2), 1 – 23.

Peeta, S. and A. Ziliaskopoulos (2001). Foundations of dynamic traffic assignment: the past, the present and the future. *Networks and Spatial Economics 1*(3), 233–265.

Ran, B. and D. Boyce (1996). *Modeling Dynamic Transportation Networks: An Intelligent Transportation System Oriented Approach*. New York: Springer-Verlag.

Ran, B., D. Boyce, and L. LeBlanc (1993). A new class of instantaneous dynamic user optimal traffic assignment models. *Operations Research 41*(1), 192–202.

Szeto, W. and H. Lo (2004). A cell-based simultaneous route and departure time choice model with elastic demand. *Transportation Research Part B 38*(7), 593–612.

Tobin, R. (1993). Notes on flow propagation constraints. Working Paper 93-10, Network Analysis Laboratory, George Mason University.

Wie, B., R. Tobin, T. Friesz, and D. Bernstein (1995). A discrete-time, nested cost operator approach to the dynamic network user equilibrium problem. *Transportation Science 29*(1), 79–92.

Wu, J., Y. Chen, and M. Florian (1998). The continuous dynamic network loading problem: a mathematical formulation and solution method. *Transportation Research Part B 32*(3), 173–187.

Zhu, D. L. and P. Marcotte (2000). On the existence of solutions to the dynamic user equilibrium problem. *Transportation Science 34*(4), 402–414.

Ziliaskopoulos, A. K. (2000). A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science 34*(1), 1–12.

Chapter 10
Dynamic Pricing and Revenue Management

An active and rapidly growing applied operations research discipline is the field known as *revenue management* (RM). The principal intent of revenue management is to extract all unused willingness to pay from consumers of differentiated services and products. Talluri and van Ryzin (2004) provide a comprehensive introduction to most aspects of the theory and practice of revenue management. For this chapter, our goal is to illustrate and solve some differential Nash games that occur in network revenue management and that provide critical information about pricing, resource allocation, and demand management to retailers and service providers.

Network RM arises in airline, railway, hotel and cruiseline service environments, where customers enter into service relationships with service providers, and those relationships may be viewed as bundles of resources. As we shall see, the relationships among service providers, customers, and resources naturally take the form of a network. Network RM also arises in the provisioning of retail businesses that sell consumer goods. By intention the scope of this book does not include stochastic models, so our discussion of revenue management in this chapter necessarily omits important details of the random aspect of demand that influences both pricing and overbooking. We hope someday to address stochastic dynamic optimization and stochastic differential games, including stochastic revenue management models that acknowledge demand uncertainty, in a revised edition.

In essentially all revenue management applications, one of the most important issues is how to model demand. Since the root tactic upon which revenue management is based is *pricing*, especially the dynamic adjustment of prices to maximize immediate or short-run revenue, the more accurate the model of demand, the more revenue one can typically generate. Another important RM tool is *demand management* through overbooking and demand rejection, a tactic that also underscores the importance of demand information. Overbooking results when a service firm accepts more reservations than its physical capacity can serve in order to hedge against cancellations and no-shows. A third revenue management tactic is *resource allocation* by which we mean the efficient use of those resources available to a service provider or retailer. Resource allocation includes the assignment of crews and vehicles to routes, network design, and other resource allocation problems familiar from applied operations research.

T.L. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research & Management Science 135,
DOI 10.1007/978-0-387-72778-3_10, © Springer Science+Business Media, LLC 2010

The following is a preview of the principal topics covered in this chapter:

Section 10.1: Dynamic Pricing with Fixed Inventories. In this section, we consider an abstract model that has some of the characteristics of revenue management, namely nonreplenishable inventories and the potential to employ dynamic pricing strategies.

Section 10.2: Revenue Management as an Evolutionary Game. In this section, we consider an abstract oligopoly and introduce market dynamics that adjust prices according the difference between current price and the moving average of price. That is, competitors learn how to set prices while simultaneously adjusting their outputs.

Section 10.3: Network Revenue Management. In this section, we allow competing Nash-firms to control pricing, resources and overbooking. The resulting generalized differential Nash game is solved heuristically.

10.1 Dynamic Pricing with Fixed Inventories

To prepare for studying dynamic pricing in a pure service setting for which outputs may not be inventoried, let us first consider a situation characterized by fixed endowments in the form of inventories that cannot be replenished since there is no production. Such a circumstance represents a kind of proto-service environment that allows us to become familiar with some of the issues that are intrinsic to dynamic pricing. Motivated by Perakis and Sood (2006), who consider a similar problem in discrete time, we consider a decision environment with the following properties:

1. At the outset, perfect information obtains about the structure of demand, the impact of price changes on demand, and the initial inventory.
2. Demand for the output of each seller is a function only of current prices, and prices are the only factor that distinguishes products from one another. That is, we assume there is a deterministic demand function faced by each seller that depends on own-period prices.
3. There is a single product, and inventory must be zero at the terminal time.
4. Sellers maximize the present value of their respective revenues by setting prices and allocating their demand to customers; they do not employ any other strategies.
5. An infinite-dimensional generalized Nash equilibrium describes the market of interest.

We will employ the notation $\pi_s(t)$ to note the price charged by seller $s \in \mathcal{S}$ at time $t \in [t_0, t_f]$, where $\mathcal{S}$ denotes the set of all sellers. We use the notation

$$\pi = (\pi_s : s \in \mathcal{S}) \in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{S}|} \tag{10.1}$$

to represent the vector of prices that are the decision variables of the model to be constructed. We use the notation

$$h_s\left[\pi(t)\right] : \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{S}|} \longrightarrow \mathcal{H}^1\left[t_0, t_f\right] \tag{10.2}$$

for the *observed demand* for the output of each seller $s \in \mathcal{S}$. We let $D_s(t)$ represent the *realized demand* experienced by seller $s \in \mathcal{S}$, and naturally define

$$D = (D_s : s \in \mathcal{S}) \in \left(L^2\left[t_0, t_f\right]\right)^{|\mathcal{S}|}$$
$$h = (h_s : s \in \mathcal{S}) \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{S}|}$$

Since realized demand must be less than or equal to observed demand, we impose the constraint

$$D_s(t) \leq h_s\left[\pi(t)\right] \quad \forall s \in \mathcal{S}, t \in \left[t_0, t_f\right]$$

or equivalently

$$D(t) \leq h\left[\pi(t)\right] \quad \forall t \in \left[t_0, t_f\right] \tag{10.3}$$

Each seller $s \in \mathcal{S}$, seeks to solve the infinite-dimensional mathematical program

$$\max J\left(\pi_s, D_s\right) = \int_{t_0}^{t_f} \exp(-\rho t)\pi_s\left(t\right) D_s(t) dt \tag{10.4}$$

subject to

$$\left(\pi_s, D_s\right) \in \Lambda_s(\pi^{-s})$$

$$D_s \leq h_s\left[\pi_s, \pi^{-s}\right] \tag{10.5}$$

$$K_s = \int_{t_0}^{t_f} D_s\left(t\right) dt \tag{10.6}$$

$$\pi_s \geq \pi_{\min} \in \Re^1_{++} \tag{10.7}$$

$$\pi_s \leq \pi_{\max} \in \Re^1_{++} \tag{10.8}$$

$$D_s \geq D_{\min} \in \Re^1_{++} \tag{10.9}$$

In the above

$$\pi^{-s} = \left(\pi_g : g \in \mathcal{S} \backslash s\right)$$

is a vector of non-own prices, while K_s is the initial endowment of inventory possessed for each seller $s \in \mathcal{S}$. Furthermore, $\pi_{\min}$ is a lower bound on prices, and $\pi_{\max}$ is the upper bound on prices. In a slight abuse of notation we take

$$\pi = \left(\pi_s, \pi^{-s}\right)$$

to be the complete column vector of prices.

10.1.1 Infinite-Dimensional Variational Inequality Formulation

It is helpful to restate the problem faced by each seller $s \in \mathcal{S}$ as the following:

$$\max J_s(\pi_s, D_s) = \int_{t_0}^{t_f} \exp(-\rho t)\pi_s(t)\, D_s(t)dt \quad \text{s.t.} \quad (\pi_s, D_s) \in \Lambda_s(\pi^{-s}) \tag{10.10}$$

where

$$\Lambda_s(\pi^{-s}) \equiv \left\{ (\pi_s, D_s) : \pi_{\min} - \pi_s \le 0, \pi_s - \pi_{\max} \le 0, \right.$$
$$\left. \int_{t_0}^{t_f} D_s(t)\, dt - K_s = 0, D_{\min} - D_s \le 0, D_s - h_s(\pi_s, \pi^{-s}) \le 0 \right\}$$

is the strategy space for seller $s \in \mathcal{S}$. Also we define

$$\Lambda = \{(\pi, D) : (\pi_s, D_s) \in \Lambda_s(\pi^{-s}) \ \ \forall s \in \mathcal{S}\} \tag{10.11}$$

We will assume each $h_s(\pi_s, \pi^{-s})$ is quasiconcave in π_s so that $\Lambda_s(\pi^{-s})$ is concave, a result formalized in the following lemma:

Lemma 10.1. *Each seller's convex strategy space. For each seller $s \in \mathcal{S}$, take $h_s(\pi_s, \pi^{-s})$ to be quasiconcave in π_s for all π^{-s}. Then the strategy space of each seller, $\Lambda_s(\pi^{-s})$, is convex in (π_s, D_s) for all $s \in \mathcal{S}$. Additionally Λ is convex.*

Proof. Note that all the constraint functions for seller $s \in \mathcal{S}$ are convex functions. In particular, let us consider the constraint

$$g_s(D_s) \equiv \int_{t_0}^{t_f} D_s(t)\, dt - K_s \le 0$$

For arbitrary points D_s^1 and D_s^2 with $\lambda \in [0, 1] \subset \Re_+^1$, we have

$$\begin{aligned} g_s\left[\lambda D_s^1(t) + (1-\lambda) D_s^2\right] &= \int_{t_0}^{t_f} \left[\lambda D_s^1(t) + (1-\lambda) D_s^2\right] dt - K_s \\ &= \lambda \int_{t_0}^{t_f} D_s^1(t)\, dt - \lambda K_s \\ &\quad + (1-\lambda) \int_{t_0}^{t_f} D_s^2(t)dt - (1-\lambda) K_s \\ &= \lambda \left\{ \int_{t_0}^{t_f} D_s^1(t)\, dt - K_s \right\} \\ &\quad + (1-\lambda) \left\{ \int_{t_0}^{t_f} D_s^2(t)dt - K_s \right\} \\ &= \lambda g_s\left(D_s^1\right) + (1-\lambda) g_s\left(D_s^2\right) \end{aligned}$$

so the constraints $g_s(D_s)$ are linear for all $s \in \mathcal{S}$ and, therefore, convex. Furthermore, since $h_s(\pi_s, \pi^{-s})$ is quasiconcave by the given, it is easy to show that

$$f_s(\pi) \equiv D_s - h_s(\pi_s, \pi^{-s}) \leq 0$$

is quasiconvex for all $s \in \mathcal{S}$. Thus, since the level sets of a quasiconvex function are convex, every $\Lambda_s(\pi^{-s})$ is convex as a set. Moreover, the convexity of the $\Lambda_s(\pi^{-s})$ for all $s \in \mathcal{S}$ assures the convexity of Λ. ■

Furthermore, we know for the function spaces stipulated that the G-derivative of the criterion functional is

$$\begin{aligned}
&\delta J_s(\pi_s, D_s; \phi_\pi, \phi_D) \\
&= \lim_{\theta \longrightarrow 0} \int_{t_0}^{t_f} \exp(-\rho t) \frac{(\pi_s + \theta\phi_\pi)(D_s + \theta\phi_D) - \pi_s D_s}{\theta} dt \\
&= \lim_{\theta \longrightarrow 0} \int_{t_0}^{t_f} \exp(-\rho t) \frac{\left(\pi_s D_s + \theta^2 \phi_\pi \phi_D + \theta D_s \phi_\pi + \theta \pi_s \phi_D\right) - \pi_s D_s}{\theta} dt \\
&= \lim_{\theta \longrightarrow 0} \int_{t_0}^{t_f} \exp(-\rho t) \frac{\theta\phi_\pi\phi_D + D_s\phi_\pi + \pi_s\phi_D}{1} dt \\
&= \int_{t_0}^{t_f} \exp(-\rho t)(D_s\phi_\pi + \pi_s\phi_D)\, dt \\
&= \int_{t_0}^{t_f} \begin{pmatrix} \exp(-\rho t) D_s \\ \exp(-\rho t)\pi_s \end{pmatrix}^T \begin{pmatrix} \phi_\pi \\ \phi_D \end{pmatrix} dt
\end{aligned} \tag{10.12}$$

Of course,

$$\delta J_s(\pi_s, D_s; \phi_\pi, \phi_D) = \int_{t_0}^{t_f} [\nabla_s J(\pi_s, D_s)]^T \phi dt \tag{10.13}$$

where

$$\phi = \begin{pmatrix} \phi_\pi \\ \phi_D \end{pmatrix}$$

Upon comparing (10.12) and (10.13), we see that

$$\nabla_s J_s(\pi_s, D_s) = \begin{pmatrix} \partial J_s / \partial \pi_s \\ \partial J_s / \partial D_s \end{pmatrix} = \begin{pmatrix} \exp(-\rho t) D_s \\ \exp(-\rho t)\pi_s \end{pmatrix}$$

The first-order condition when $\left(\pi_s^*, D_s^*\right) \in \Lambda_s(\pi^{-s})$ is a solution is

$$\delta J\left(\pi_s^*, D_s^*; \phi_\pi, \phi_D\right) \leq 0$$

for all feasible directions

$$\phi_\pi = \pi_s - \pi_s^*$$
$$\phi_D = D_s - D_s^*$$

The above is easily restated in the more familiar form

$$\left\langle \nabla_s J_s\left(\pi_s^*, D_s^*\right), \begin{pmatrix} \pi_s - \pi_s^* \\ D_s - D_s^* \end{pmatrix} \right\rangle \leq 0$$

or equivalently as

$$\int_{t_0}^{t_f} \left[\frac{\partial J_s\left(\pi_s^*, D_s^*\right)}{\partial \pi_s}\left(\pi_s - \pi_s^*\right) + \frac{\partial J_s\left(\pi_s^*, D_s^*\right)}{\partial D_s}\left(D_s - D_s^*\right) \right] dt \leq 0$$
$$\forall (\pi_s, D_s) \in \Lambda_s(\pi^{-s})$$

In light of our knowledge of the gradient of the criterion J_s, this last variational inequality may be stated as

$$\int_{t_0}^{t_f} \exp(-\rho t)\left[D_s^* \cdot \left(\pi_s - \pi_s^*\right) + \pi_s^* \cdot \left(D_s - D_s^*\right)\right] dt \leq 0 \quad \forall (\pi_s, D_s) \in \Lambda_s(\pi^{-s}) \tag{10.14}$$

Statement (10.14) will not only be a necessary condition but also a sufficient condition if $J_s(\pi_s, D_s)$ is pseudoconcave on $\Lambda_s(\pi^{-s})$. The pseudoconcavity of each seller's criterion will allow us to establish an equivalent variational inequality formulation of the generalized Nash equilibrium among sellers described by (10.10). To that end we state and prove the following result:

Lemma 10.2. *Seller's criterion is pseudoconcave. When the best response of each seller $s \in \mathcal{S}$ is constrained to be strictly positive and bounded away from the origin, each criterion $J_s(\pi_s, D_s)$ is pseudoconcave on $\Lambda_s(\pi^{-s})$.*

Proof. For the criterion $J_s(\pi_s, D_s)$ to be pseudoconcave on $\Lambda_s(\pi^{-s})$, we must show that

$$\left\langle \nabla_s J_s\left(\pi_s^2, D_s^2\right), \begin{pmatrix} \pi_s^1 - \pi_s^2 \\ D_s^1 - D_s^2 \end{pmatrix} \right\rangle \leq 0 \Longrightarrow J_s\left(\pi_s^2, D_s^2\right) \geq J_s\left(\pi_s^1, D_s^1\right) \tag{10.15}$$

for $\left(\pi_s^1, D_s^1\right), \left(\pi_s^2, D_s^2\right) \in \Lambda_s(\pi^{-s})$. Property (10.15) is assured if $R_s(\pi_s, D_s) = \pi_s D_s$ is pseudoconcave at each instant of time $t \in \left[t_0, t_f\right]$. By Theorem 9 of Ferland (1972), we know that

$$Z_s = (-1) \cdot R_s(\pi_s, D_s)$$

is pseudoconvex (and hence $R_s(\pi_s, D_s)$ is pseudoconcave), if the following matrix

$$M_s = \begin{pmatrix} 0 & \dfrac{\partial Z_s}{\partial \pi_s} & \dfrac{\partial Z_s}{\partial D_s} \\ \dfrac{\partial Z_s}{\partial \pi_s} & \dfrac{\partial^2 Z_s}{\partial \pi_s^2} & \dfrac{\partial^2 Z_s}{\partial \pi_s \partial D_s} \\ \dfrac{\partial Z_s}{\partial D_s} & \dfrac{\partial^2 Z_s}{\partial D_s \partial \pi_s} & \dfrac{\partial^2 Z_s}{\partial D_s^2} \end{pmatrix} = \begin{pmatrix} 0 & -D_s & -\pi_s \\ -D_s & 0 & -1 \\ -\pi_s & -1 & 0 \end{pmatrix}$$

has a determinant that is strictly negative. We note that

$$\det M_s = -2 D_s \pi_s < 0$$

since by the given each seller $s \in \mathcal{S}$ employs a strictly positive solution bounded away from the origin. ■

We now present a key result:

Theorem 10.1. *The generalized Nash game among seller's captured by (10.10) for all $s \in \mathcal{S}$, is equivalent to the infinite-dimensional variational inequality*

$$\sum_{s \in \mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t) \left[D_s^* \cdot \left(\pi_s - \pi_s^* \right) + \pi_s^* \cdot \left(D_s - D_s^* \right) \right] dt \leq 0 \quad \forall (\pi, D) \in \Lambda \tag{10.16}$$

when each demand function is concave in own price.

Proof. Clearly we have established above that (10.14) is a necessary condition for the Nash equilibrium of interest. Summing (10.14) over $s \in \mathcal{S}$ yields (10.16). It remains for us to establish sufficiency. However, for each $s \in \mathcal{S}$, the pseudoconcavity of $R_s(\pi_s, D_s)$ assures (10.14) is a sufficient condition for (10.10). If given variational inequality (10.16), by selecting $D_s = D_s^*$ and $\pi_s = \pi_s^*$ for all $s \neq r$, the minimum principle is recovered for each seller $r \in [1, |\mathcal{S}|]$. Equivalency is thereby established. ■

10.1.2 Restatement of the Isoperimetric Constraints

Each seller $s \in \mathcal{S}$ will face the constraint

$$\int_{t_0}^{t_f} D_s(t)\, dt = K_s \tag{10.17}$$

which has the form of an isoperimetric constraint. As such, (10.17) may be restated as the following two-point boundary-value problem:

$$\frac{dy_s}{dt} = D_s \tag{10.18}$$

$$y_s(t_0) = 0 \tag{10.19}$$

$$y(t_f) = K_s \tag{10.20}$$

where y_s is a newly introduced state variable.

10.1.3 Differential Variational Inequality Formulation

By exploiting the two-point boundary-value problem (10.18), (10.19), and (10.20), the infinite-dimensional variational inequality (10.16) may be given the following form: find $(\pi^*, D^*) \in \Omega$ such that

$$\sum_{s \in \mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t) \left[D_s^* \cdot \left(\pi_s - \pi_s^* \right) + \pi_s^* \cdot \left(D_s - D_s^* \right) \right] dt \geq 0 \quad \forall (\pi, D) \in \Omega \tag{10.21}$$

where

$$\Omega = \{(\pi, D) : (\pi_s, D_s) \in \Omega_s(\pi^{-s}) \ \ \forall s \in \mathcal{S}\}$$

and

$$\Omega_s(\pi^{-s}) = \Big\{ (\pi_s, D_s) : \frac{dy_s}{dt} = D_s,\ y_s(t_0) = 0,\ y(t_f) = K_s,\ \pi_{\min} - \pi_s \leq 0,$$
$$\pi_s - \pi_{\max} \leq 0,\ D_{\min} - D_s \leq 0,\ D_s - h_s(\pi_s, \pi^{-s}) \leq 0 \Big\}$$

The reader should recognize that (10.21) is a continuous-time differential quasivariational inequality.

10.1.4 Numerical Example

Because we are dealing with a differential quasivariational inequality, algorithms that enjoyed proofs of convergence for mere differential variational inequalities will not ncessarily converge when applied to (10.21). So we now proceed to apply a gap-function algorithm heuristically; if convergence is acheived, the result will be a generalized differential Nash equilibrium. With these remarks in mind, consider the following simple example of three firms with fixed inventories and both upper and low bounds on their prices and the demands they fulfill. We elect to solve this problem using time discretization and a finite-dimensional gap function. The three sellers' initial endowments are

$$K_1 = 3000$$
$$K_2 = 2000$$
$$K_3 = 2500$$

The initial and terminal times are $t_0 = 0$ and $t_f = 5$, respectively. Observed demand for the output of seller $s = 1, 2$, or 3 is

$$h_s[\pi(t)] = a_s(t) - b_s(t)\pi_s(t) + \sum_{i \neq s} c_i(t)\pi_i(t)$$

where

$$\begin{aligned} a_1 &= 300 \\ a_2 &= 200 \\ a_3 &= 300 \\ b_s(t) &= 2 - 0.2t \quad s = 1, 2, 3 \\ c_s(t) &= 0.1t + 0.5 \quad s = 1, 2, 3 \end{aligned}$$

Upper and lower bounds on their prices and demands are as follow:

$$D_{\min} = 0 \qquad \pi_{\min} = \begin{pmatrix} 50 \\ 35 \\ 45 \end{pmatrix} \qquad \pi_{\max} = \begin{pmatrix} 300 \\ 300 \\ 300 \end{pmatrix}$$

We will solve this example using the D-gap function method. Recall that, once a differential gap function has been formed, it is used to create a nonlinear program whose solution is also a solution of the differential variational inequality of interest.

This example is expressible as a differential variational inequality having the control vector

$$u = \begin{pmatrix} \pi \\ D \end{pmatrix}$$

The set of feasible controls is

$$U = \{u : \pi_{\min} - \pi_s \leq 0 \;\; \pi_s - \pi_{\max} \leq 0 \;\; D_s \geq 0 \;\; D_s - h_s(\pi_s, \pi^{-s}) \;\; s = 1, 2, 3\}$$

The following definitions also apply:

$$F = \begin{pmatrix} D \\ \pi \end{pmatrix} \qquad f = \frac{dy}{dt} = D$$

$$\Psi\left[y\left(t_f\right), t_f\right] = \begin{pmatrix} y_1(10) \\ y_2(10) \\ y_3(10) \end{pmatrix} = \begin{pmatrix} 3000 \\ 2000 \\ 2500 \end{pmatrix}$$

Of course, $DVI(F, f, \Psi, U, x_0, t_0, t_f)$ denotes the differential variational inequality of interest.

In this example, we employ a D-gap function of the form

$$\psi_{\alpha\beta} = \varphi_\alpha(u) - \varphi_\beta(u)$$

where

$$\varphi_\alpha(u) = \max < F(y,u,t), v-u > -\frac{\alpha}{2}\|v-u\|^2$$

$$\varphi_\beta(u) = \max < F(y,u,t), v-u > -\frac{\beta}{2}\|v-u\|^2$$

Therefore

$$\psi_{\alpha\beta} =< F(y,u,t), v_\beta(u)-v_\alpha(u) > -\frac{\alpha}{2}\|v_\alpha(u)-u\|^2 + \frac{\beta}{2}\|v_\beta(u)-u\|^2$$

where

$$v_\alpha(u) = P_U\left[u-\frac{1}{\alpha}F(y,u,t)\right] = P_U\left[\begin{pmatrix}\pi\\ D\end{pmatrix} - \frac{1}{10}\begin{pmatrix}D\\ \pi\end{pmatrix}\right]$$

$$v_\beta(u) = P_U\left[u-\frac{1}{\beta}F(y,u,t)\right] = P_U\left[\begin{pmatrix}\pi\\ D\end{pmatrix} - \frac{1}{10.2}\begin{pmatrix}D\\ \pi\end{pmatrix}\right]$$

Then the gradient information we need is

$$\nabla\psi_{\alpha\beta} = \frac{\partial F(y,u,t)}{\partial u}\left[v_\beta(u)-v_\alpha(u)\right]+\alpha\left[v_\alpha(u)-u\right]-\beta\left[v_\beta(u)-u\right]+\lambda\frac{\partial f(y,u,t)}{\partial u}$$

We employ the constant step size $\theta_k = 0.5$ and solve the above differential variational inequality using the D-gap function, after discretizing to create a finite dimensional problem. The discrete time approximation employed involves $N = 1000$ equal time steps. GAMS/PATHNLP supported by Matlab was used to solve this problem. We set $\alpha = 10$ and $\beta = 10.2$, where evidently $0 < \alpha < \beta$. The algorithm was terminated after 10000 iterations with a gap of about .03. The results generated using the aforementioned computing scheme are summarized in this table:

Iteration k	Gap $\psi_{\alpha\beta}(u^k)$
1	16.1861
2	0.1335
3	0.1333
⋮	⋮
500	0.1132
1000	0.0922
2000	0.0412
5000	0.0368
8000	0.0312
10000	0.0283

Figures 10.1 and 10.2 provide depictions of realized demand and price trajectories.

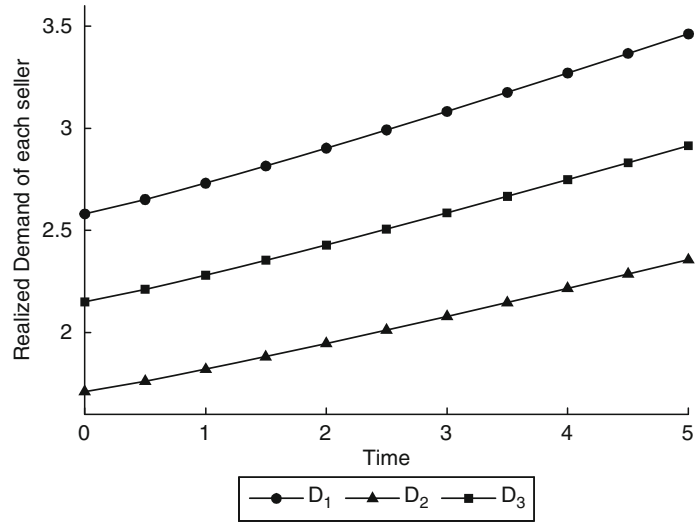


Fig. 10.1 Realized demand trajectories of sellers

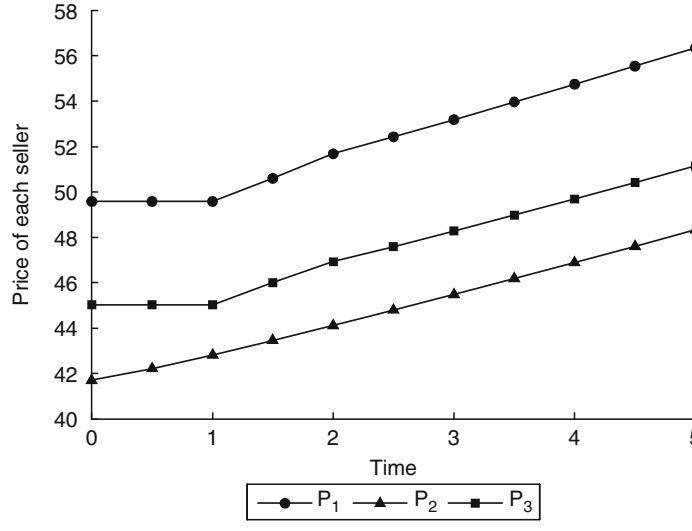


Fig. 10.2 Price trajectories of sellers

10.2 Revenue Management as an Evolutionary Game

Although demand theoretically devolves from utility maximization by individuals, an actual demand curve and its parameters are seldom available for most markets. We want now to model the dynamics of demand as a differential equation based on an evolutionary game theory (EGT) perspective. To that end, we express a dynamic

nonzero sum evolutionary game among service providers as a differential variational inequality. The service providers of interest will be viewed as having fixed upper bounds on output derived from capacity constraints on available resources.

The service providers of interest in this section are in oligopolistic game-theoretic competition according to a learning process that can be likened to the learning mechanisms considered in evolutionary game theory and for which price changes are proportional to their signed excursion from a market clearing price. We stress that in this model firms are setting prices for their services while simultaneously determining the levels of demand they will serve. This is unusual in that, typically, firms in oligopolistic competition are modeled in microeconomic theory as setting either prices or output flows, but not both. The joint adjustment of prices and outputs is modeled here by comparing current price to the price that would have cleared the market for the demand that has most recently been served. However, the service providers are unable to make this comparison until the current round of play is completed as knowledge of the total demand served by all competitors is required.

Kachani et al. (2004) put forward a revenue management model for service providers to address such joint pricing and demand learning in an oligopolistic setting with fixed capacity constraints. The model they consider assumes the demand faced by a service provider is a linear function of its price and other competitors' prices, although the impact of a change in price on demand in one period does not automatically propagate to later time periods. In our presentation below, we allow this impact to propagate to all subsequent time periods. Furthermore, we consider only a single class of customers, so-called bargain-hunting buyers searching for personal or, to a limited extent, business services or products at the most competitive prices; these buyers are willing to sacrifice some convenience for the sake of a lower price. Because the services and products are assumed to be homogeneous, ties between two sellers offering the same price are broken arbitrarily. In other words, the consumer has no concept of brand preference in the decision environment we consider.

10.2.1 Assumptions and Notation

Assume that a set of service providers are competing in an oligopolistic setting, each with the objective of maximizing their revenue. These service providers have very high fixed costs compared to their relatively low variable or operating costs. Therefore, each provider focuses only on maximizing its own revenue. Moreover, each firm provides a set of services, each of which is homogeneous. For example, the difference between an economy class seat on Southwest Airlines and an economy class seat on Jet Blue is indiscernible by customers; the only differences that the customers perceive are the prices charged by the different service providers. Furthermore, every service provider can set the price for each of its services. The price charged for each service in one time period will affect the demand for that service in the next time period. The price a service provider charges is compared to the moving

average price over all competitors. The rate of service provision by each company has an upper bound. A provider must, therefore, choose prices that create demand for its services, maximize revenue, and ensure capacities are not exceeded.

We denote the set of revenue managing firms by $\mathcal{F}$, each of whom is providing a set of services $\mathcal{S}$. Continuous time is denoted by the scalar $t \in \Re^1_+$, while t_0 is the initial time and $t_f \in \Re^1_{++}$ the finite terminal time so that $t \in [t_0, t_f] \subset \Re^1_+$. Each firm $f \in \mathcal{F}$ controls prices

$$\pi_i^f \in L^2[t_0, t_f]$$

corresponding to each service type $i \in \mathcal{S}$. The control vectors of individual firms are

$$\pi^f \in \left(L^2[t_0, t_f]\right)^{|\mathcal{S}|} \quad \forall f \in \mathcal{F}$$

which are then concatenated to form the complete vector of controls

$$\pi \in \left(L^2[t_0, t_f]\right)^{|\mathcal{S}| \times |\mathcal{F}|},$$

We also let

$$D_i^f(\pi, t) : \left(L^2[t_0, t_f]\right)^{|\mathcal{S}| \times |\mathcal{F}|} \times \Re^1_+ \longrightarrow \mathcal{H}^1[t_0, t_f]$$

denote the demand for service $i \in \mathcal{S}$ of firm $f \in \mathcal{F}$ and define the vector of all such demands for firm f to be

$$D^f \in \left(\mathcal{H}^1[t_0, t_f]\right)^{|\mathcal{S}|}$$

All such demands for service $i \in \mathcal{S}$ of firm $f \in \mathcal{F}$ are denoted by

$$D \in \left(\mathcal{H}^1[t_0, t_f]\right)^{|\mathcal{S}| \times |\mathcal{F}|}$$

We will use the notation

$$D^{-f} = \left(D_i^g : i \in \mathcal{S}, g \in \mathcal{F} \setminus \{f\}\right)$$

for the vector of service levels provided by the competitors of firm $f \in \mathcal{F}$.

10.2.2 Demand Dynamics

In evolutionary game theory the notion of comparing a moving average to the current state is used to develop ordinary differential equations describing learning processes; see Fudenberg and Levine (1999). As we have already stressed, the services provided by different agents are homogeneous; hence, customers' decisions depend only on prices. The demand for the service offerings of firm $f \in \mathcal{F}$ evolve

according to the following dynamics:

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\tilde{\pi}_i - \pi_i^f\right) \quad \forall i \in \mathcal{S},\ \ f \in \mathcal{F} \tag{10.22}$$

$$D_i^f\,(t_0) = K_{i,0}^f \qquad \forall i \in \mathcal{S},\ \ f \in \mathcal{F} \tag{10.23}$$

where $\tilde{\pi}_i$ is the moving average price for service $i \in \mathcal{S}$ given by

$$\tilde{\pi}_i\,(t) = \frac{1}{|\mathcal{F}|\,(t - t_0)} \int_{t_0}^{t} \sum_{g \in \mathcal{F}} \pi_i^g\,(\tau) d\tau \quad \forall i \in \mathcal{S}$$

while $K_{i,0}^f \in \Re_{++}^1$ and $\eta_i^f \in \Re_{++}^1$ are exogenous parameters for each $i \in \mathcal{S}$ and $f \in \mathcal{F}$. The firms set the parameter η_i^f by analyzing past demand data and the elasticity of demand with respect to price. The demand for service type i provided by firm f changes over time in accordance with the excess between the firm's price and the moving average of all agents' prices for the particular service. The coefficient η_i^f influences how quickly demand reacts to price changes for each firm f and service type i. Some providers may specialize in certain services and may be able to adjust more quickly than their competitors. These dynamics are reminiscent of so-called replicator dynamics which are used in evolutionary games; see Hofbauer and Sigmund (1998). The rate of growth of demand, may be viewed as the rate of growth of firm f with respect to service type i. This growth follows the "basic tenet of Darwinism" in that it may be interpreted as the difference between the fitness (price charged) of the firm providing a given service and the moving average of the fitness of all firms providing that same service.

10.2.3 Constraints

There are positive upper and lower bounds, reflecting market regulations or known customer behavior, on service prices charged by firms. Thus, we write

$$\pi_{\min,i}^f \leq \pi_i^f \leq \pi_{\max,i}^f \quad \forall i \in \mathcal{S},\ \ f \in \mathcal{F}$$

where the $\pi_{\min,i}^f \in \Re_{++}^1$ and $\pi_{\max,i}^f \in \Re_{++}^1$ are known constants. Similarly, there will be a lower bound on the demand for services of each type by each firm as negative demand levels are meaningless; that is,

$$D_i^f \geq 0 \ \ \forall i \in \mathcal{S},\ \ f \in \mathcal{F}$$

Let $\mathcal{R}$ be the set of resources that the firms can utilize to provide services, while $|\mathcal{R}|$ is the cardinality of $\mathcal{R}$. Define an arbitrary element of the incidence matrix

$$A = (a_{lm})$$

by

$$a_{lm} = \begin{cases} 1 & \text{if resource } l \text{ is used by the service type } m \\ 0 & \text{otherwise} \end{cases}$$

Joint resource-constraints for each firm f are also imposed and take the form

$$0 \leq A \cdot D_i^f \leq C_i^f \quad \forall i \in \mathcal{S}, \ \ f \in \mathcal{F} \tag{10.24}$$

where C_i^f denotes the resources available for use by firm $f \in \mathcal{F}$ in providing service i.

10.2.4 The Firm's Optimal Control Problem

Since revenue management firms have negligible variable costs and high fixed costs, each firm's objective is to maximize revenue which in turn ensures that profit is maximized. The instantaneous revenue for firm f is

$$\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f$$

Consequently, each firm $f \in \mathcal{F}$ faces the following extremal problem: with the π^{-f} as exogenous inputs, solve the following optimal control problem:

$$\max_{\pi^f} J_f(\pi^f, \pi^{-f}, t) = \int_{t_0}^{t_f} e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) dt - e^{-\rho t_0} \Psi_0^f \tag{10.25}$$

subject to

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\tilde{\pi}_i - \pi_i^f \right) \quad \forall i \in \mathcal{S}, \ \ f \in \mathcal{F} \tag{10.26}$$

$$D_i^f(t_0) = K_{i,0}^f \quad \forall i \in \mathcal{S} \tag{10.27}$$

$$\pi_{\min,i}^f \leq \pi_i^f \leq \pi_{\max,i}^f \quad \forall i \in \mathcal{S} \tag{10.28}$$

$$0 \leq A \cdot D_i^f \leq C_i^f \quad \forall i \in \mathcal{S} \tag{10.29}$$

where Ψ_0^f is the fixed cost of production for firm f which is later dropped from the formulation, ρ is the nominal discount rate compounded continuously, and

$$\int_{t_0}^{t_f} e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) dt$$

is the net present value of revenue. We may restate these dynamics for all $f \in \mathcal{F}$ as

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}|\,(t - t_0)} - \pi_i^f \right) \quad \forall i \in \mathcal{S} \tag{10.30}$$

$$\frac{dy_i}{dt} = \sum_{g \in \mathcal{F}} \pi_i^g \quad \forall i \in \mathcal{S} \tag{10.31}$$

$$D_i^f\,(t_0) = K_{i,0}^f \quad \forall i \in \mathcal{S} \tag{10.32}$$

$$y_i\,(t_0) = 0 \quad \forall i \in \mathcal{S} \tag{10.33}$$

As a consequence we may rewrite the optimal control problem of firm $f \in \mathcal{F}$ as

$$\max_{\pi^f} J_f(\pi^f, \pi^{-f}, t) = \int_{t_0}^{t_f} e^{-\rho t} \left(\sum_{i \in \mathcal{S}} \pi_i^f \cdot D_i^f \right) dt \tag{10.34}$$

subject to

$$\frac{dD_i^f}{dt} = \eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}|\,(t - t_0)} - \pi_i^f \right) \quad \forall i \in \mathcal{S} \tag{10.35}$$

$$\frac{dy_i}{dt} = \sum\nolimits_{g \in \mathcal{F}} \pi_i^g \quad \forall i \in \mathcal{S} \tag{10.36}$$

$$D_i^f\,(t_0) = K_{i,0}^f \quad \forall i \in \mathcal{S} \tag{10.37}$$

$$y_i\,(t_0) = 0 \quad \forall i \in \mathcal{S} \tag{10.38}$$

$$\pi_{\min,i}^f \le \pi_i^f \le \pi_{\max,i}^f \quad \forall i \in \mathcal{S} \tag{10.39}$$

$$-D_i^f \le 0 \quad \forall i \in \mathcal{S} \tag{10.40}$$

$$A \cdot D_i^f - C_i^f \le 0 \quad \forall i \in \mathcal{S} \tag{10.41}$$

In condensed notation, this generalized differential Nash game can be expressed as: with the π^{-f} as exogenous inputs, each firm must compute π^{f*} that solves the following optimal control problem:

$$\begin{aligned} &\max\ J_f(\pi^f, \pi^{-f}, t) \\ &\text{s.t.}\quad \pi^f \in \Omega_f\left(\pi^{-f}\right) \end{aligned} \tag{10.42}$$

for all $f \in \mathcal{F}$ where

$$\Omega_f\left(\pi^{-f}\right) = \left\{ \pi^f : (10.35), (10.36), (10.37), (10.38), (10.39), (10.40), \text{and } (10.41) \text{ hold} \right\}$$

10.2.5 Differential Quasivariational Inequality Formulation

Each service provider is a Nash agent that knows and employs the current instantaneous values of the decision variables of other firms to make its own noncooperative decisions. Clearly, (10.42) defines a set of coupled optimal control problems, one for each firm $f \in \mathcal{F}$. It is useful to note that each instance of (10.42) is an optimal control problem with a fixed terminal time and a fixed terminal state. Its Hamiltonian is

$$H_f\left(D^f, y, \pi^f, \lambda^f, \sigma, \alpha^f, \beta^f; \pi^{-f}; t\right) \equiv e^{-\rho t}\left(\sum_{i\in\mathcal{S}} \pi_i^f \cdot D_i^f\right) + \Phi_f\left(\pi^f; D^f; \lambda^f; \sigma^f; \alpha^f; \beta^f; \pi^{-f}\right)$$

where

$$\begin{aligned}\Phi_f\left(D^f, y, \pi^f, \lambda^f, \sigma, \alpha^f, \beta^f; \pi^{-f}; t\right) \equiv & \sum_{i\in\mathcal{S}} \lambda_i^f \left[\eta_i^f \cdot \left(\frac{y_i}{|\mathcal{F}|\,(t-t_0)} - \pi_i^f\right)\right] \\ & + \sum_{i\in\mathcal{S}} \sigma_i \left(\pi_i^f + \sum_{g\in\mathcal{F}\setminus f} \pi_i^g\right) \\ & + \sum_{i\in\mathcal{S}} \alpha_i^f \left(-D_i^f\right) \\ & + \sum_{i\in\mathcal{S}} \beta_i^f \left(A \cdot D_i^f - C_i^f\right) \qquad (10.43)\end{aligned}$$

while $\lambda_i^f \in \mathcal{H}^1\left[t_0, t_f\right]$ is the adjoint variable for the demand dynamics associated with the firm $f \in \mathcal{F}$ and service type $i \in \mathcal{S}$, while $\lambda^f \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{S}|}$. Furthermore, $\sigma_i \in \mathcal{H}^1\left[t_0, t_f\right]$ is the adjoint variable for the dynamics describing the dummy state variables $y_i \in \mathcal{H}^1\left[t_0, t_f\right]$ for all $i \in \mathcal{S}$, while $\sigma \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{S}|}$ and $y \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{S}|}$; and of course $D_i^f \in \mathcal{H}^1\left[t_0, t_f\right]$ and $D \in \left(\mathcal{H}^1\left[t_0, t_f\right]\right)^{|\mathcal{F}|\cdot|\mathcal{S}|}$. Additionally, for all $i \in \mathcal{S}$ and $f \in \mathcal{F}$, the $\alpha_i^f \in \Re_+^1$ and $\beta_i^f \in \Re_+^1$ are dual variables for the state space constraints (10.40) and (10.41), respectively.

We assume that the game arising from the coupled optimal control problems (10.42) is such that all the operations performed previously and below are well defined and the necessary conditions we have introduced in previous chapters are also sufficient. Therefore, the maximum principle tells us that an optimal solution to (10.42) is the tuple $\left(D^{*f}, y^*, \pi^{*f}, \lambda^{*f}, \sigma^*, \alpha^{*f}, \beta^{*f}\right)$ that solves the nonlinear program

$$\max H_f \quad \text{s.t.} \quad \pi_{\min}^f \leq \pi^f \leq \pi_{\max}^f$$

for every firm $f \in \mathcal{F}$ for every instant of time $t \in \left[t_0, t_f\right]$ where

$$\pi^f_{\min} = \left\{\pi^f_{\min,i} : i \in \mathcal{S}\right\}$$
$$\pi^f_{\max} = \left\{\pi^f_{\max,i} : i \in \mathcal{S}\right\}$$

That is, any optimal solution must satisfy at each time $t \in \left[t_0, t_f\right]$

$$\pi^{f*} = \arg\left\{\max_{\pi^f_{\min} \le \pi^f \le \pi^f_{\max}} H_f\left(D^f, y, \pi^f, \lambda^f, \sigma, \alpha^f, \beta^f; \pi^{-f}; t\right)\right\} \tag{10.44}$$

which in turn is equivalent to

$$\left[\nabla_{\pi^f} H^*_f\right]^T \left(\pi^f - \pi^{*f}\right) dt \le 0 \tag{10.45}$$

when

$$\begin{pmatrix} \pi^f_{\min} \\ \pi^f_{\min} \end{pmatrix} \le \begin{pmatrix} \pi^{*f} \\ \pi^f \end{pmatrix} \le \begin{pmatrix} \pi^f_{\max} \\ \pi^f_{\max} \end{pmatrix} \tag{10.46}$$

where

$$H^*_f \equiv e^{-\rho t}\left(\sum_{i \in \mathcal{S}} \pi_i^{f*} \cdot D_i^{f*}\right) + \Phi^*_f \tag{10.47}$$

and

$$\Phi^*_f = \Phi_f\left(D^{f*}, y^*, \pi^{f*}, \lambda^{f*}, \sigma^*, \alpha^{f*}, \beta^{f*}; \pi^{-f*}; t\right) \tag{10.48}$$

Furthermore, the relevant adjoint dynamics include

$$\frac{\partial H_f}{\partial D^f_i} = (-1)\frac{d\lambda_i^{f*}}{dt} \tag{10.49}$$

Due to absence of terminal time constraints, transversality requires

$$\lambda^{f*}\left(t_f\right) = \gamma^T \frac{\partial \Gamma\left[D^{f*}\left(t_f\right), t_f\right]}{\partial D^{f*}\left(t_f\right)} = 0 \tag{10.50}$$

which, when taken together with the state dynamics and the dynamics for y, gives rise to a two-point boundary-value problem.

With the preceding background, we are now in a position to create a variational inequality for noncooperative competition among the firms. In particular, we have immediately from (10.45) the following differential quasivariational inequality that has dynamic generalized Nash equilibria as solutions:

$$\int_{t_0}^{t_f} \left[\sum_{s\in\mathcal{S}} \sum_{f\in\mathcal{F}} \frac{\partial H_f^*}{\partial \pi_i^f} \left(\pi_i^f - \pi_i^{f*} \right) \right] dt \leq 0 \quad \forall \pi \in \Lambda(\pi) \equiv \prod_{f\in\mathcal{F}} \Lambda_f \left(\pi^{-f} \right) \tag{10.51}$$

where H_f^* is defined by (10.47) and (10.48) and we recall that

$$\Lambda_f \left(\pi^{-f} \right) = \left\{ \pi^f \in \Omega_f \left(\pi^{-f} \right) : (10.49) \text{ and } (10.50) \text{ hold} \right\}$$

This differential quasivariational inequality is a convenient way of expressing the generalized Nash game that is our present interest. The reader will find it instructive to itemize all the assumptions implicit in 10.51. We leave as an exercise for the reader a formal demonstration that a solution of (10.51) is a solution of the underlying generalized differential Nash game (10.42).

10.2.6 Numerical Example

Let us consider an abstract revenue management scenario wherein five service providers are involved in oligopolistic competition. Each of these firms offers a set of four services and compete for the market demand of these services. The minimal prices are

$$\pi_{\min,i}^f = 0 \quad \forall i \in \mathcal{S}, f \in \mathcal{F}$$

The remaining parameters used for this example are given in the tables provided below. In particular, the demand sensitivity parameters are:

$\eta_i^f : 10^{-3} \times$

Service type, i	1	2	3	4
Firm 1	10	8	12	9
Firm 2	11	7	10	15
Firm 3	20	12	20	20
Firm 4	15	10	15	18
Firm 5	18	6	20	20

The initial demands are:

$K_{i,0}^f :$

Service type, i	1	2	3	4
Firm 1	10	17.5	22.5	30
Firm 2	9.5	16.5	20	31
Firm 3	10.5	17	25	28.5
Firm 4	11	19	24	31
Firm 5	10.5	18	23	30.5

The resource upper bounds are:

C_i^f :

Service type, i	1	2	3	4
Firm 1	50	35	25	10
Firm 2	30	25	20	12.5
Firm 3	51	37.5	26	10.5
Firm 4	45	37.5	28	11
Firm 5	42	35	27.5	10.5

The price upper bounds are:

$\pi_{\max,i}^f$:

Service type, i	1	2	3	4
Firm 1	8.5	13.5	18	20.5
Firm 2	7.5	10.8	18.5	21
Firm 3	8.6	11.7	15.5	20.1
Firm 4	9	11	16.5	20.5
Firm 5	9.2	11.6	17.5	21.8

The revenue management and pricing problem that the firms face is to continuously set the prices of their four services for $t \in [t_0, t_f]$ where $t_0 = 0$ and $t_f = 5$. The nominal discount rate is $\rho = 0.05$, compounded continuously. Because the we are dealing with differential quasivariational inequalities, the algorithms presented in Chapter 6 do not enjoy convergence proofs. Hence, we elect to heuristically apply a fixed-point algorithm, solving its subproblems by gradient projection after time discretization. If convergence occurs, a generalized differential Nash equilibrium will have been achieved. The fixed-point stopping tolerance is set at

$$\varepsilon = 0.01$$

Additionally, we choose

$$\alpha = \frac{1}{k}$$

where k is the fixed-point major iteration counter and α is the arbitrary positive coefficient of the fixed-point formulation.

We forgo the detailed symbolic statement of this example and, instead, provide numerical results in graphical form for the solution after 113 fixed-point iterations. Figure 10.3 shows the price trajectories for the services set by the firms as well as the moving average of price for each service type. Figure 10.4 depicts how demands for the services of each firm change over time in response to the prices set by the firms. The instantaneous revenues generated over time by the firms are plotted in

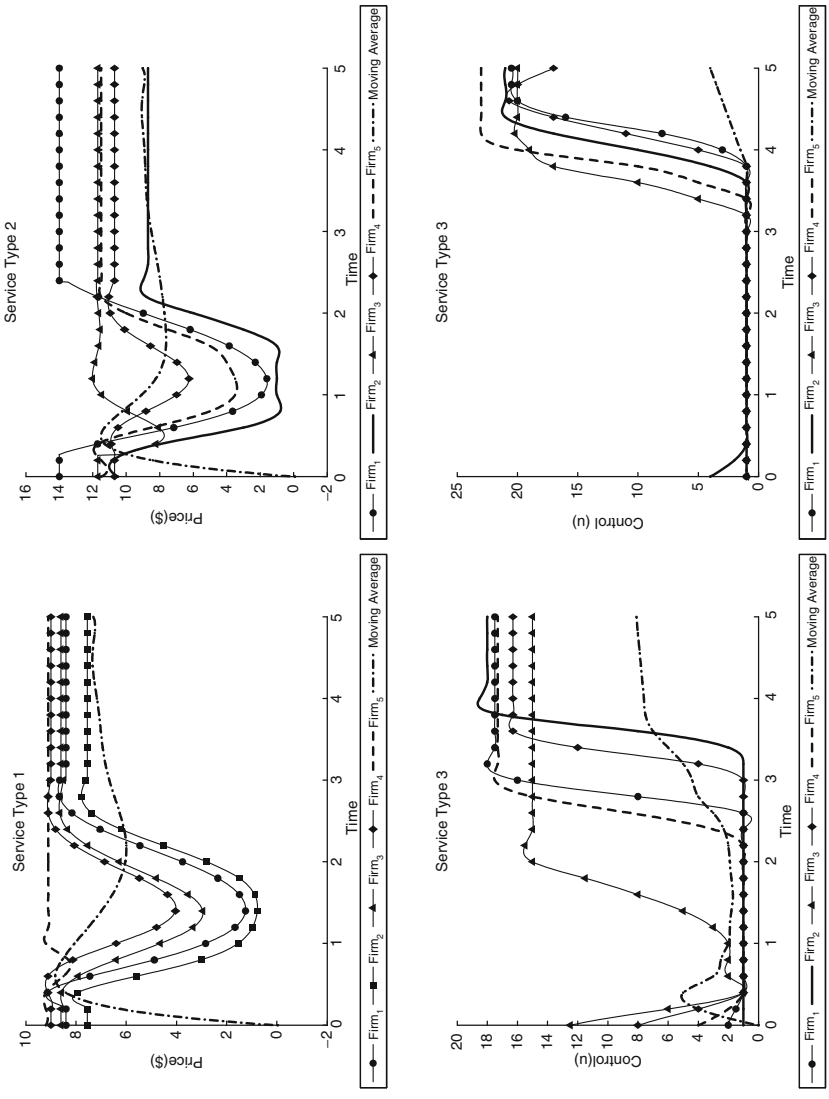


Fig. 10.3 Instantaneous and moving average price trajectories

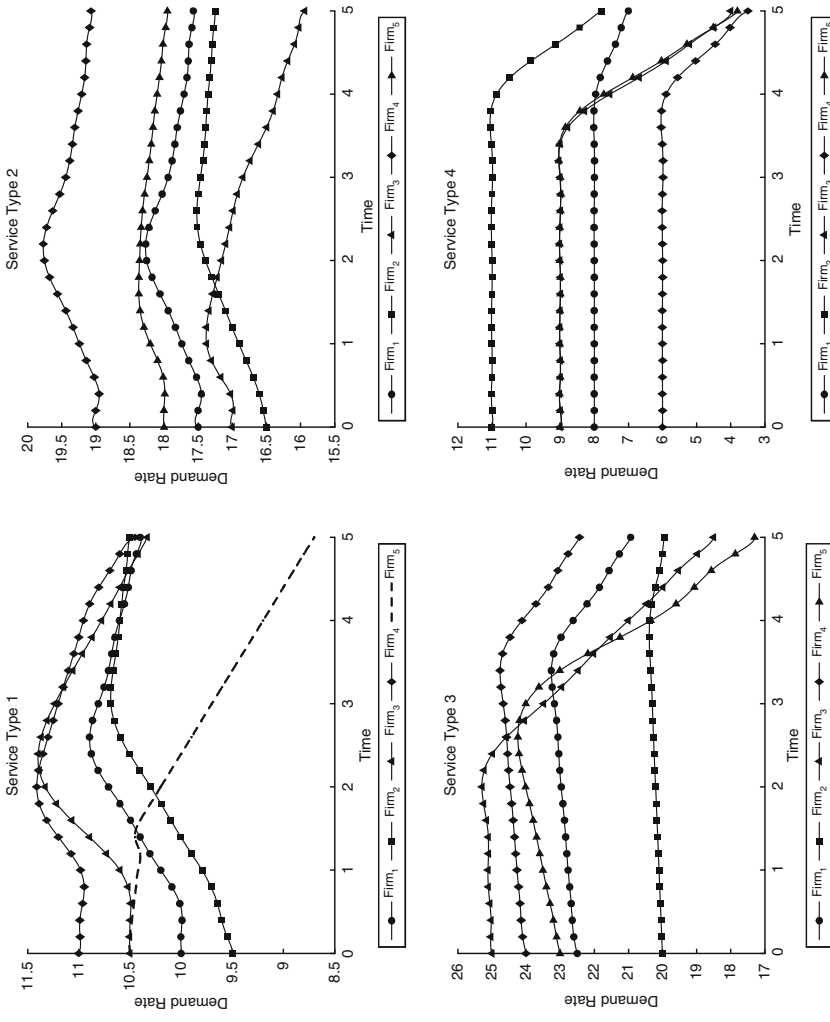


Fig. 10.4 Firms' demand dynamics for different service types (grouped by service types)

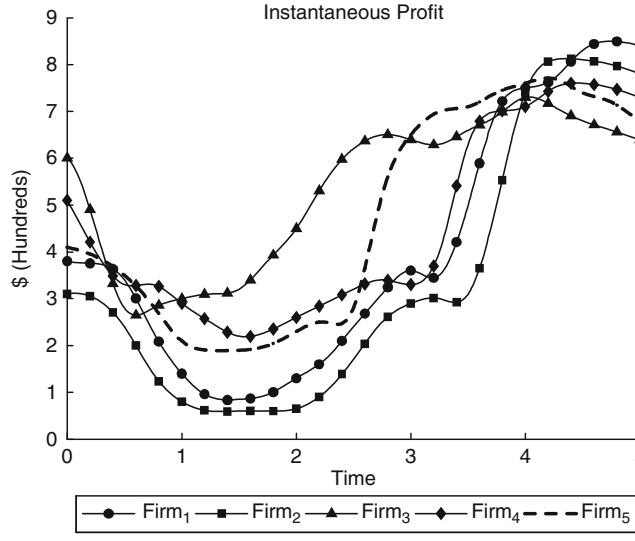


Fig. 10.5 Instantaneous revenue generated by the firms

Figure 10.5. The net present vaues of revenue generated by the firms at the end of planning horizon are given in the following table:

Firm	Net Present Values
Firm 1	8586
Firm 2	7268
Firm 3	11128
Firm 4	9484
Firm 5	10026

The parameters of this numerical example depict a scenario wherein service type 1 is a low-valued service and service type 4 is a high-valued service. Firm 2 is a discounter with modest capacity, while firm 5 is a more expensive service provider. Demand for firm 3 is more sensitive to price than is the case for the other firms. A fascinating behavior is observed for all the firms in their price-setting mechanisms: even though they initially set different prices for a service, toward the end of the planning horizon, all firms start behaving similarly and their prices converge. Furthermore, possibly singular controls are observed for all firms. This only affirms the difficult nature of pricing in a competitive, nonlinear environment. Even the slightest deviation from the optimal trajectories may cause the firms to suffer dramatic performance degradation. Finally, the net present values of cumulative revenue show that the discounting firm 2, which offers deep discounts, cannot effectively exploit its low price structure and lags behind the others in the long run. Additionally, we see that firm 3 benefits most in the competition, even though it is not a discounter.

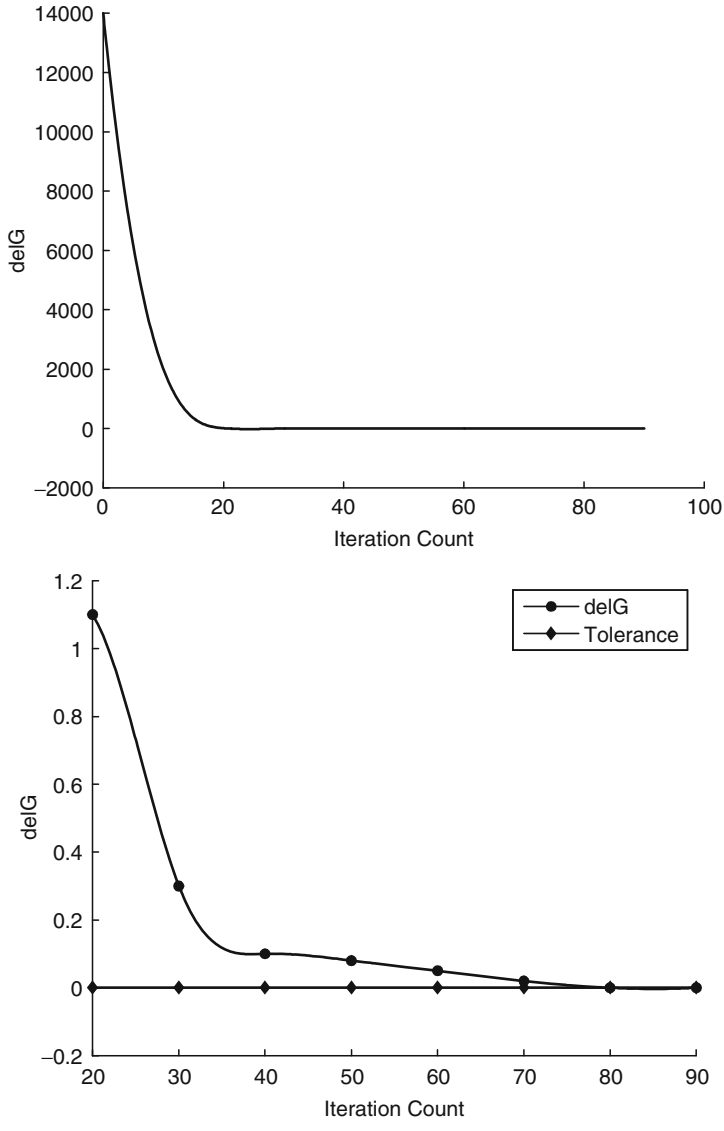


Fig. 10.6 (Up) Relative change of controls from one iteration to the next (Δ_k) vs. the iteration counter (k), (Down) the zoomed-in view

As mentioned earlier, the fixed-point algorithm converged after 113 iterations for this numerical example. In Figure 10.6 the relative change from one iteration to the next, expressed as

$$\Delta_k = \left\| \pi^k - \pi^{k-1} \right\| \leq \varepsilon$$

is plotted against the iteration counter k. It is worth noting that for this particular example even though $\Delta_1 = 3.4 \times 10^3$, the next several iterations very rapidly decrease Δ_k. The run time for this example is less than 3 minutes using a generic desktop computer with single Intel Pentium processors and 1 GB RAM. The computer code for the continuous-time fixed-point algorithm was written in MatLab 6.5 and calls a gradient projection subroutine implemented in discrete time.

10.3 Network Revenue Management

In this section we consider service firms that provide differentiated, nonsubstitutable services, set prices for their services, may decline booking requests at any given time, face cancellations and no-shows (with full or partial refunds), and have finite supplies of resources. The demand is assumed to be known with certainty, a severe assumption mandated by the deterministic focus of this book. Stochastic extensions of the model reported here are found in Mookherjee and Friesz (2008). Each service firm has to decide both how to allocate its resources (quantity-based RM) and how to set prices for its services (price-based RM) as it seeks to maximize its revenue. Such a decision environment differs from that faced by discount airlines only in the absence of demand uncertainty.

10.3.1 Discrete-Time Notation

We consider an oligopoly of abstract service providers. Each firm provides a set of services (products). Each network service is to be viewed as a bundle of resources sold with certain terms of purchase and restrictions at a given price. These services are nonsubstitutable and differentiated. All firms have finite resource capacities. The booking period is taken to be $[0, L]$ which is discretized in N time segments. For this model, at the beginning of discrete period $t \in [0, N]$, firms set service prices and quantities for sale in that period. The notation we will use for sets and matrices is given in the following table:

Set/Matrix	Definition
$\mathcal{F}$	set of firms
$\mathcal{S}$	set of services each firm provides
$\mathcal{C}$	set of resources that firms use to provide services
$\mathcal{C}_i$	set of resources that firms use to provide service $i \in \mathcal{S}$
$\mathcal{S}_j$	set of services that utilize resource $j \in \mathcal{C}$
$\mathcal{A}$	resource-service incidence matrix
$\lvert\mathcal{C}\rvert$	cardinality of $\mathcal{C}$, with analogous definitions for other sets

The notation for model parameters is the following:

$$
\begin{aligned}
L &= \text{end time of the booking period} \\
0 \le t_n \le t_N &= L \text{ constrains } n\text{th period of the booking horizon} \\
\Delta &= \frac{L}{N-1} \text{ defines inter-period time step } t_{n+1} - t_n \\
p^f_{i,\min} &\in \Re^1_{++} \text{ is the minimum price charged by firm } f \text{ for service } i \in \mathcal{S} \\
p^f_{i,\min} &\in \Re^1_{++} \text{ is the maximum price charged by firm } f \text{ for service } i \in \mathcal{S} \\
x_{\min} &\in \Re^1_{++} \text{ strictly positive minimum allowed level of service} \\
K^f_i &\in \Re^1_{++} \text{ capacity of firm } f \in \mathcal{F} \text{ for resource type } j \in \mathcal{C} \\
\rho^f_t &\in \Re^1_{++} \text{ cancellation rate for firm } f \in \mathcal{C} \text{ at end of period } t
\end{aligned}
$$

The notation for variables and functions, for period $t \in [0, N-1]$, is the following:

$$
\begin{aligned}
p^f_{i,t} &= \text{price for service } i \in \mathcal{S} \text{ charged by firm } f \in \mathcal{F} \\
x^f_{j,t} &= \text{ resource allocation by firm } f \in \mathcal{F} \text{ of type } j \in \mathcal{C} \\
D^f_{i,t}(p_{i,t}) &= \text{demand realized by firm } f \in \mathcal{F} \text{ for service } i \in \mathcal{S} \\
R^f_{j,t}(\rho^f_t \cdot x^f_t) &= \text{refund by firm } f \in \mathcal{F} \text{ for cancelling resource } j \in \mathcal{C}
\end{aligned}
$$

Additionally $\Psi^f_N(x^f_N, K^f)$ is the denial of service cost for firm $f \in \mathcal{F}$. The vector of prices for service i provided by firm f is

$$p^f_i = \left(p^f_{i,t} : t \in [0, N]\right)$$

We will also need to work with the vectors

$$
\begin{aligned}
p^f_t &= \left(p^f_{i,t} : i \in \mathcal{S}\right) \\
p^f &= \left(p^f_t : t \in [0, N]\right) \\
p_t &= \left(p^f_t : f \in \mathcal{F}\right)
\end{aligned}
$$

The pricing decision variables of firm f's competitors for period t are denoted by the vector

$$p^{-f}_t = \left(p^g_t : g \in \mathcal{F} \backslash f\right),$$

The state variables for firm f are the vectors of cumulative allocations of resources

$$x^f_t = \left(x^f_{j,t} : j \in \mathcal{C}\right)$$

for period t.

The network we are interested in has $|\mathcal{C}|$ resources and the firm provides $|\mathcal{S}|$ different services. As already noted, each network product is a combination of a bundle of the $|\mathcal{C}|$ resources sold with certain terms of purchase and restrictions at a given price. The resource-service *incidence matrix*, $\mathcal{A} = \left[a_{ij}\right]$ is a $|\mathcal{C}| \times |\mathcal{S}|$ matrix where

$$a_{ij} = \begin{cases} 1 & \text{if resource } i \text{ is used for service } j \\ 0 & \text{if resource } i \text{ is not used for service } j \end{cases}$$

Thus, the jth column of $\mathcal{A}$, denoted by $\mathcal{A}_j$, is the *incidence vector* for service j; while the ith row, denoted by $\mathcal{A}^i$, has unity in column j provided service j utilizes resource i. Note that there may be multiple identical columns of $\mathcal{A}$ if there are multiple ways of selling a given bundle of resources, although each could have different revenue values and different demand patterns; see Talluri and van Ryzin (2004) for a more detailed discussion of this and related subtleties.

10.3.2 Demand Functions

For our demand model we assume that the customers make their purchasing decisions based on the current period's price only; hence, demand for any period depends only on the price vector for that period only. Firm f's realized demand, or *bookings*, for service i during time period t, when the prevailing market price is p_t, will be denoted as $D_{i,t}^f(p_{i,t})$. The following are an exhaustive set of regularity assumptions for demand suggested by Mookherjee and Friesz (2008). We do not actually employ all of them in the analyses that follow; yet, since each has a behavioral foundation, it in convenient to assume all demand functions considered are regular in the sense of the below definition:

Definition 10.1. *Demand regularity. For any firm $f \in \mathcal{F}$ and service type $i \in \mathcal{S}$, the demand function $D_{i,t}^f(p_{i,t})$ is said to be regular if it displays the following properties for all periods $t \in [0, N]$:*

1. *Each price $p_{i,t}^f$ is defined on a range $\left[p_{i,\min}^f, p_{i,\max}^f\right]$, where $p_{i,\max}^f \in \Re_{++}^1$ is the maximum admissible value of p_i^f and $p_{i,\min}^f \in \Re_{++}^1$ is the minimum admissible value of p_i^f, while $D_i^f(p_{i,t})\Big|_{p_{i,t}^f = p_{i,\max}^f} = 0$;*
2. *$D_{i,t}^f(p_{i,t})$ depends only on the current period $t \in [0, N]$ prices charged by firm f and its competitors for service type i;*
3. *$D_{i,t}^f(p_{i,t})$ is continuous, bounded, and differentiable for all $p_{i,t} \in [p_{i,\min}, p_{i,\max}]$ where $p_{i,t} = \left\{p_{i,t}^g : g \in \mathcal{F}\right\}$*
4. $\dfrac{\partial D_{i,t}^f(p_{i,t})}{\partial p_{i,t}^f} < 0$

5. *The own-price elasticity of* $D_{i,t}^{f}\left(p_{i,t}\right)$, *defined as* $e_{i,t}^{f} \equiv -\left(\frac{\partial D_{i,t}^{f}}{\partial p_{i,t}^{f}} \Big/ \frac{D_{i,t}^{f}}{p_{i,t}^{f}}\right)$ *is nondecreasing in* $p_{i,t}^{f}$; *that is,* $\frac{\partial e_{i,t}^{f}}{\partial p_{i,t}^{f}} \geq 0$;
6. $\frac{\partial D_{i}^{f}\left(p_{i,t}\right)}{\partial p_{i,t}^{g}} > 0$ *for all* $g \neq f$;
7. $\frac{\partial e_{i,t}^{f}}{\partial p_{i,t}^{g}} \leq 0$ *for all* $g \neq f$; *and*
8. $\frac{\partial e_{i,t}^{f}}{\partial p_{i,t}^{f}} + \sum_{g \neq f} \frac{\partial e_{i,t}^{f}}{\partial p_{i,t}^{g}} \geq 0$ *for all* $i \in \mathcal{S},\ f \in \mathcal{F}$.

It is instructive to give qualitative descriptions for each of the separate conditions above. In particular, items 1 and 3 impose bounds that arise from regulations and policy. Item 2 stipulates that consumers of services do not make direct intertemporal price comparisons. Items 4 and 5 indicate that $D_{i,t}^{f}(p_{i,t})$ is downward sloping in firm f's own price and has nondecreasing price elasticity relative to $p_{i,t}^{f}$. Item 6 states that when any other firm increases its price for service i, there is a corresponding increment of firm f's demand for the service i. Item 7 further requires that an increase in firm g's price for service i not only increases firm f's demand, but also decreases firm f's price elasticity. Item 8 stipulates that the local price effect of a price change dominates the cross price effect on the local price elasticity.

It is easy to verify that most of the commonly used demand functions satisfy the restrictions set forth above in our notion of demand regularity. In particular, the following demand functions are regular in the sense of Definition 10.1:

1. *Linear*
$$D_{i,t}^{f}\left(p_{i,t}\right) = \rho_{i,t}^{f} - \sigma_{i,t}^{f} \cdot p_{i,t}^{f} + \sum_{g \in \mathcal{F} \setminus f} \gamma_{i,t}^{g} \cdot p_{i,t}^{g}$$
where $\rho_{i,t}^{f}, \sigma_{i,t}^{f}, \gamma_{i,t}^{f} \in \Re_{++}^{1}$ for all $f \in \mathcal{F}, i \in \mathcal{S}$, and $0 \leq t \leq N$.
2. *Logit*
$$D_{i,t}^{f}\left(p_{i,t}\right) = \frac{a_{i,t}^{f} \exp\left(-b_{i,t}^{f} \cdot p_{i,t}^{f}\right)}{\theta_{i} + \sum_{g \in \mathcal{F}} a_{j,t}^{f} \exp\left(-b_{j,t}^{f} \cdot p_{j,t}^{f}\right)}$$
where $a_{i,t}^{f}, b_{i,t}^{f}, \theta_{i} \in \Re_{++}^{1}$ for all $f \in \mathcal{F}, i \in \mathcal{S}$, and $0 \leq t \leq N$.
3. *Cobb-Douglas*
$$D_{i,t}^{f}\left(p_{i,t}\right) = a_{i}^{f}\left(p_{i,t}^{f}\right)^{-\beta_{i}^{f}} \prod_{g \in \mathcal{F} \setminus i}\left(p_{i,t}^{g}\right)^{\beta_{i}^{fg}}$$
where $a_{i}^{f} > 0, \beta_{i}^{f} > 1, \beta_{i}^{fg} > 0$ for all $f \in \mathcal{F}, i \in \mathcal{S}$, and $0 \leq t \leq N$.

10.3.3 Denial-of-Service Costs and Refunds

In our model, deterministic demand and so-called show demand are identical; thus there is no meaning in such a setting to overbooking. However, we do allow service providers to cancel scheduled resources. In particular, cancellations are assumed to occur at the rate ρ_t^f for firm $f \in \mathcal{F}$ and discrete-time period $t \in [0, N-1]$. Such cancellations require refunds expressed as

$$R_t^f \left(\rho_t^f \cdot x_t^f \right) \tag{10.52}$$

for firm $f \in \mathcal{F}$ in period $t \in [0, N-1]$. Refunds $R_t^f(\cdot)$ should monotonically increase with x_t^f and ρ_t^f, and decrease with timė t; such qualitative behavior reflects the potential for cancellation fees to increase as the end of the booking period is approached. Denial of service must necessarily involve loss of goodwill on the part of customers toward service providers. These denial-of-service costs are calculated at the end of the booking period and involve the comparison of resources delivered to actual capacity. Denial-of-service costs are expressed as

$$\Psi_N^f \left(x_N^f, K^f \right) \tag{10.53}$$

for firm $f \in \mathcal{F}$, where of course

$$K^f = \left(K_j^f : j \in \mathcal{C} \right) \tag{10.54}$$

is the vector of actual capacities.

10.3.4 Firms' Extremal Problem

With the rival firms' prices p_t^{-f} taken as exogenous to the discrete-time optimal control problem of firm $f \in \mathcal{F}$ and yet endogenous to the overall model, firm $f \in \mathcal{F}$ computes its prices p_t^f and allocation of resources x_t^f in order to maximize net revenue generated throughout the booking period; this behavior we express as

$$\max_{p^f} J \left(p^f ; p^{-f} \right) = -\Psi_N^f \left(x_N^f, K^f \right) - \sum_{t=0}^{N-1} R_t^f \left(\rho_t^f \cdot x_t^f \right) + \sum_{t=0}^{N-1} p_t^f \cdot D_t^f \left(p_t \right) \tag{10.55}$$

subject to

$$x_{t+1}^f = x_t^f + \mathcal{A} \cdot D_t^f \left(p_t \right) - \rho_t^f \cdot x_t^f \quad t = 0, \ldots, N-1 \tag{10.56}$$

$$x_0^f = 0 \tag{10.57}$$

$$p_{\min}^f \leq p_t^f \leq p_{\max}^f \qquad t = 0, \ldots, N-1 \tag{10.58}$$

$$D_t^f\left(p_t\right) \geq x_{\min} \qquad t = 0, \ldots, N-1 \tag{10.59}$$

The first two terms on the righthand side of (10.55) are the denial-of-service costs and total refunds, respectively. These are subtracted from total revenue generated to give net revenue generated in the booking period. As in a typical service industry, there is no salvage value of unsold resources at the end of the horizon. Constraints (10.56) are definitional dynamics that describe the net rate of resource commitment. Of course (10.57) is an initial condition that states no resources are committed at the start of the booking period. Service prices are bounded from above and below in (10.58). Constraint (10.59) serves to bound realized demand away from zero; without this constraint, it is possible for a service provider to offer no service in one or more periods yet to set prices, which would be implausible. Furthermore, it is clear that the above defines a generalized differential Nash equilibrium; we therefore expect the associated variational inequality to be a differential quasivariational inequality.

Since the firms' optimal control problems are coupled, we are dealing with a dynamic Nash game. Observe that service provider f's resource allocations x_t^f impacts his/her own revenue but not that of any of his/her competitors whereas service price does. Hence a firm only needs information on his/her competitors' pricing policies and not information on their allocations, as is appropriate since the latter would be unrealistic in practice. Let us now form the discrete-time Hamiltonian

$$\begin{aligned} H_f \equiv H_f\left(p^f; \lambda^f; p^{-f}; t\right) = \sum_{t=0}^{N-1} \Big[& p_t^f \cdot D_t^f\left(p_t\right) - R_t^f\left(\rho_t^f \cdot x_t^f\right) \\ & + \left(\lambda_{t+1}^f\right)^T \cdot \left(x_t^f + \mathcal{A} \cdot D_t^f\left(p_t\right) - \rho_t^f \cdot x_t^f\right) \Big] \end{aligned} \tag{10.60}$$

where λ^f is the vector of adjoint variables such that

$$\lambda_t^f = \left(\lambda_{j,t}^f : j \in \mathcal{C}\right)$$

$$\lambda^f = \left(\lambda_t^f : t \in [1, N]\right)$$

The adjoint variables λ^f may be interpreted as the *shadow prices* of resources. From the maximum principle, at each time period $t \in [0, N-1]$ firm f seeks to solve the following static optimization problem

$$\max_{p_t^f} H_{f,t} \tag{10.61}$$

subject to

$$p_{\min}^{f} \leq p_{t}^{f} \leq p_{\max}^{f} \tag{10.62}$$

$$D_{t}^{f}\left(p_{t}\right) \geq x_{\min} \tag{10.63}$$

The adjoint dynamics arising from the optimal control problem of each firm $f \in \mathcal{F}$ are

$$\lambda_{j,t+1}^{f} - \lambda_{j,t}^{f} = (-1)\frac{\partial H_{f}}{\partial x_{j,t}^{f}} \text{ for all } j \in \mathcal{C} \tag{10.64}$$

Also the transversality condition for each firm $f \in \mathcal{F}$ is

$$\lambda_{j,N}^{f} = (-1)\frac{\partial \Psi_{N}^{f}\left(x_{N}^{f}, K^{f}\right)}{\partial x_{j,N}^{f}} \text{ for all } j \in \mathcal{C} \tag{10.65}$$

A necessary condition for p_{t}^{f*} to be a solution of the best response problem can be expressed as the following variational inequality: find $p^{f*} \in \mathcal{K}_{f}$ such that

$$\left[\nabla_{p^{f}} H_{f}\left(p^{f*}; \lambda^{f}; p^{-f}; t\right)\right]^{T} \cdot \left(p^{f} - p^{f*}\right) \leq 0 \tag{10.66}$$

for all $p_{i,t}^{f} \in \Lambda_{f}$ where

$$\mathcal{K}_{f}\left(p^{-f}\right) = \left\{p^{f} : (10.56), (10.57), (10.58), (10.59), (10.64), \text{ and } (10.65)\right\} \tag{10.67}$$

10.3.5 Market Equilibrium Problem as a Quasivariational Inequality

With the preceding background, we can now formulate the market equilibrium problem as a variational inequality. We combine the variational inequalities (10.66) for each firm $f \in \mathcal{F}$ and time $t \in [0, N-1]$. We define the following feasible space for all service providers:

$$\mathcal{K}(p) = \prod_{f \in \mathcal{F}} \mathcal{K}_{f}\left(p^{-f}\right) \tag{10.68}$$

The discrete-time differential quasivariational inequality of interest seeks to find $p^{*} \in \mathcal{K}(p^{*})$ such that

$$\begin{pmatrix} \nabla_{p^1} H_1\left(p^{1*}; \lambda^{1*}; p^{-1*}\right) \\ \vdots \\ \nabla_{p^{|\mathcal{F}|}} H_{|\mathcal{F}|}\left(p^{|\mathcal{F}|*}; \lambda^{|\mathcal{F}|*}; p^{-|\mathcal{F}|*}\right) \end{pmatrix}^T \cdot \begin{pmatrix} p^1 - p^{1*} \\ \vdots \\ p^{|\mathcal{F}|} - p^{|\mathcal{F}|*} \end{pmatrix} \leq 0 \qquad (10.69)$$

for all $p \in \mathcal{K}(p)$. We leave the formal demonstration, including the imposition of regularity conditions, that solutions of (10.69) are generalized differential Nash equilibria as an exercise for the reader.

10.3.6 Numerical Example

For the example that follows, we have elected to heuristically apply a fixed-point algorithm, meant for problems that are merely variational inequalities, to our present quasivariational inequality; if convergence occurs, we will have found a discrete-time generalized differential Nash equilibrium. Such an algorithm relies on our ability to re-express the variational inequality (10.69) as a fixed-point problem, a task that has been illustrated repeatedly in previous chapters and which we, therefore, do not dwell on here. The numerical example we present here is motivated by the notion of airline revenue management; the network employed is a simplified version of that found in Mookherjee and Friesz (2008).

In particular, we consider the six-node and eight-leg network of Figure 10.7. Two firms (f and g) are competing over this network. To keep the exposition simple, we assume that each firm uses the same network; of course this may be relaxed to a more general setting where each firm has its own service network that is coupled to its competitors via congestion or other externalities. We consider nine different

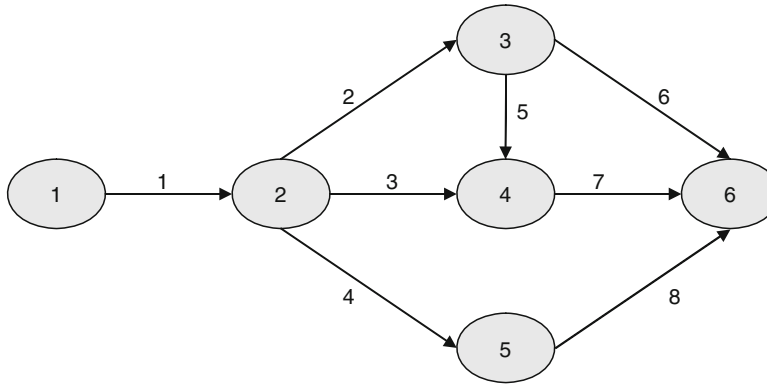


Fig. 10.7 Six-node, eight-Leg airline network

paths (services) for this network that connect five different origin-destination pairs as shown in the following table:

Service ID	O-D	Itinerary	Service ID	O-D	Itinerary
1	$(1,2)$	$1-2$	6	$(1,6)$	$1-2-4-6$
2	$(1,3)$	$1-2-3$	7	$(1,6)$	$1-2-5-6$
3	$(1,4)$	$1-2-4$	8	$(1,6)$	$1-2-3-4-6$
4	$(1,5)$	$1-2-5$	9	$(1,4)$	$1-2-3-4$
5	$(1,6)$	$1-2-3-6$			

The relevant sets of firms and services are

$$\mathcal{F} = \{1,2\}$$
$$\mathcal{S} = \{1,2,3,4\}$$

We assume that the relevant demand functions for offered services have the form

$$d_i^f(p_i,t) = (\alpha_i^f(t) - \beta_i^f(t) \cdot p_i^f(t) + \gamma_i^{f,g}(t) \cdot p_i^g(t)) \cdot |\sin(wt)|$$

for all $f \in \mathcal{F}, i \in \mathcal{S}$, and $t \in [t_0, t_1]$; the parameters $\alpha_i^f(t), \beta_i^f(t), \gamma_i^{f,g}(t) \in \Re_{++}^1$ are specified in the following table:

Service	1	2	3	4	5	6	7	8	9
$\alpha_i^{f_1}$	60	50	40	60	45	50	50	30	100
$\alpha_i^{f_2}$	25	50	60	75	35	50	30	60	45
$\beta_i^{f_1}$	.3	.3	.2	.5	.25	.2	.15	.25	.15
$\beta_i^{f_2}$	.2	.25	.3	.15	.45	.1	.2	.5	.3
$\gamma_i^{f_1,f_2}$	.15	.2	.1	.15	.2	.1	.2	.2	.15
$\gamma_i^{f_2,f_1}$	.1	.1	.15	.2	.15	.2	.1	.1	.1

In addition, w is assumed to be 20. The booking period runs from clock time $t_0 = 0$ to clock time $t_f = 20$. The bounds on service prices are:

Service	1	2	3	4	5	6	7	8	9
$p_{i,\min}^f$	50	50	75	25	30	20	50	50	75
$p_{i,\max}^f$	180	150	200	160	120	80	180	150	200

Cancellations are assumed to occur at the rates $1-\rho_j^t$ for all $j \in \mathcal{C}$ and discrete-time periods $t \in [0, N-1]$. Such cancellations require refunds expressed as

$$R_{jt}^f\left(\rho_{jt}^f \cdot x_t^f\right) = 15(1-\rho_j^t)x_j$$

The parameters ρ_j for each $j \in \mathcal{C}$ are indicated below:

Leg	1	2	3	4	5	6	7	8
ρ_j	.9	.8	.9	.85	.9	.7	.9	.8

Denial of service must necessarily involve loss of goodwill on the part of customers toward service providers. These denial-of-service costs are calculated at the end of the booking period and involve the comparison of resources delivered to actual capacity. Denial-of-service costs are expressed as

$$\Psi_N^f\left(x_N^f, K^f\right) = 300(\max(x_j - K_j^f, 0))^2$$

where

$$K^f = \left(K_j^f : j \in \mathcal{C}\right)$$

for each firm $f \in \mathcal{F}$. Each service provider has different, yet mostly comparable, capacities on each of the 8 legs; these are given in the table below:

Legs	1	2	3	4	5	6	7	8
Firm 1	600	400	400	250	250	300	300	250
Firm 2	500	250	300	250	300	300	300	500

Only bidirectional pricing is considered. For the sake of brevity, we forgo the detailed symbolic statement of this example and, instead, provide numerical results in graphical form in Figure 10.8 and Figure 10.9.

10.4 Exercises

1. Develop a model of dynamic monopolistic pricing with fixed inventories; express your model as a differential variational inequality. Develop a pricing decision rule by analyzing the necessary conditions for your formulation.
2. Establish existence under plausible regularity conditions for the dynamic pricing model with fixed inventories presented in Section 6.7.
3. Establish existence under plausible regularity conditions for the network revenue management model of Section 10.3.
4. Discuss the advantages and disadvantages of discrete versus continuous-time formulations of dynamic pricing with fixed inventories. In particular, discuss the question of uniqueness for both formulations.

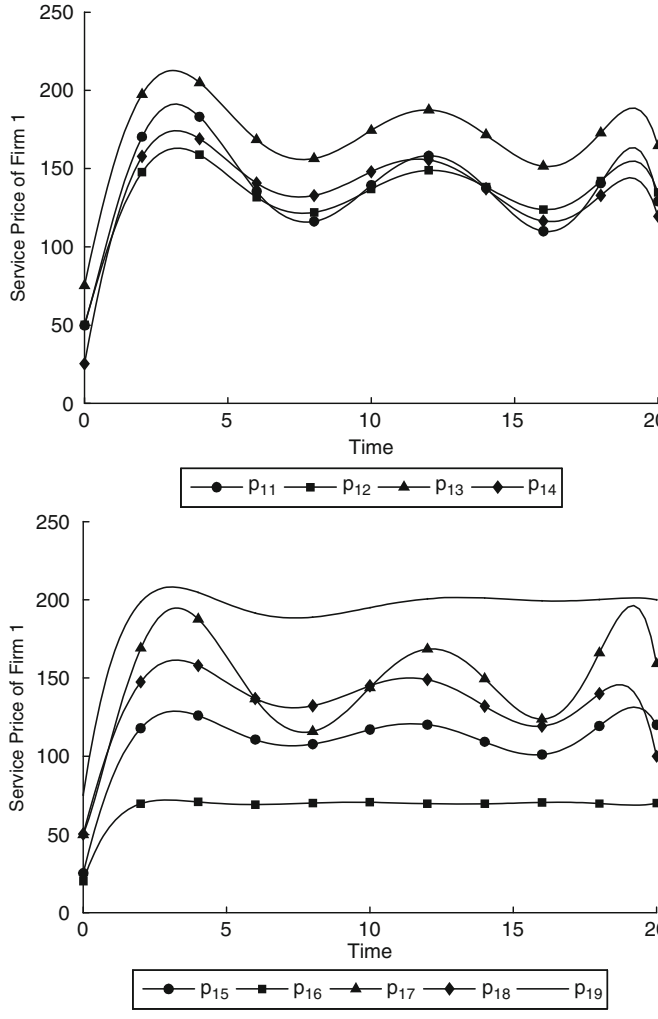


Fig. 10.8 Service price trajectory of firm 1

5. Consider the generalized differential Nash game of Section 10.2. Formally show, for appropriate regularity conditions, that a solution of (10.42) is in fact a generalized differential Nash equilibrium.
6. Consider the discrete time network revenue management model of Section 10.3. Formally show, for appropriate regularity conditions, that a solution of (10.69) is in fact a generalized differential Nash equilibrium.
7. After reading Chapter 10, create a Stackelberg leader-follower pricing model, and describe its domain of application. Use as your model of followers the evolutionary game pricing model of Section 10.2.

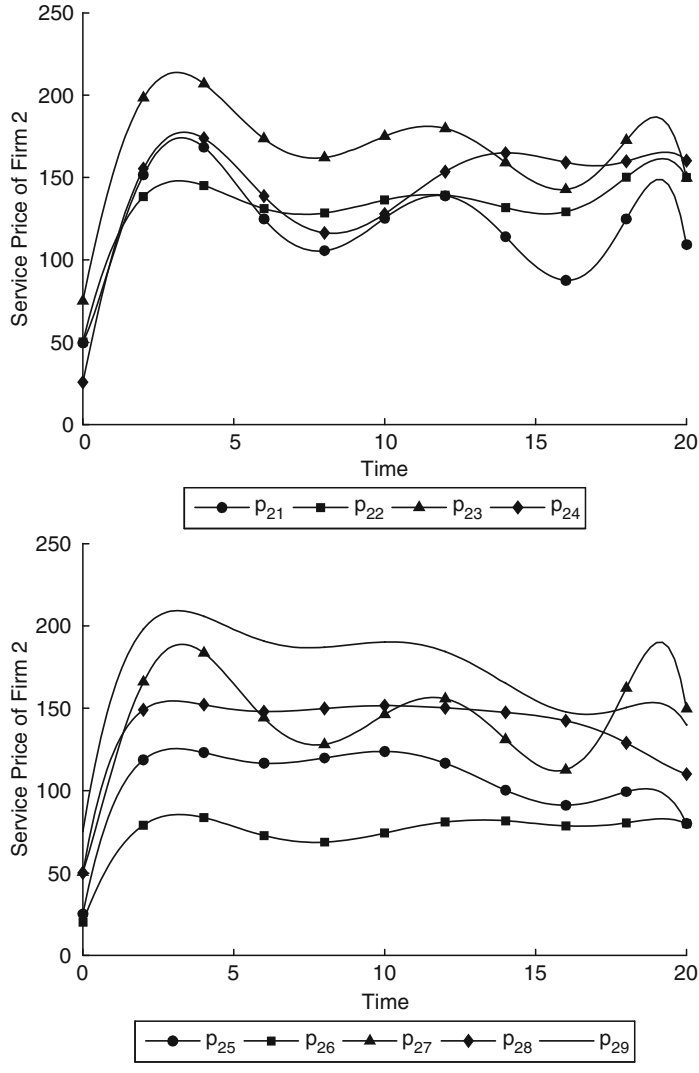


Fig. 10.9 Service price trajectory of firm 2

List of References Cited and Additional Reading

Ferland, J. A. (1972). Mathematical programming with quasi-convex objective functions. *Mathematical Programming 3*, 296–301.

Fudenberg, D. and D. K. Levine (1999). *The Theory of Learning in Games* (2nd ed.). Cambridge, MA: MIT Press.

Hofbauer, J. and K. Sigmund (1998). *Evolutionary Games and Replicator Dynamics*. Cambridge: Cambridge University Press.

Kachani, S., G. Perakis, and C. Simon (2004). A transient model for joint pricing and demand learning under competition. Presented in the 4th Annual INFORMS Revenue Management and Pricing Section Conference, MIT, Cambridge.

Mookherjee, R. and T. Friesz (2008). Pricing, allocation, and overbooking in dynamic service network competition when demand is uncertain. *Production and Operations Management 14*(4), 1–20.

Perakis, G. and A. Sood (2006). Competitive multiperiod pricing for perishable products: a robust optimization approach. *Mathematical Programming B 107*, 295–335.

Talluri, K. T. and G. J. van Ryzin (2004). *The Theory and Practice of Revenue Management*. New York: Springer-Verlag.

Index

A

Adjoint equation, 74–76, 108, 119, 121, 124, 127, 169, 274, 297, 301, 326, 330, 335–339, 360, 363, 365, 401–403, 430, 433, 442

Adjoint variable, 72, 76, 108, 116, 117, 125, 128–130, 132, 133, 142, 168, 173, 179, 216, 279, 286, 287, 306, 326, 327, 330, 334, 335, 361, 364, 366, 381, 402, 403, 440, 452, 473, 486

Admissible control, 104, 105, 170, 271, 277, 287, 384, 430

Admissible set, 436

Algorithm, 1, 20, 31, 37, 78, 130, 133, 134, 148, 149, 154, 191–210, 214, 239–242, 246–248, 253, 255–260, 282–289, 298, 303–307, 309–311, 350, 373, 385, 408, 432, 434, 439, 445–449, 452, 464, 466, 476, 480, 481, 488

Arc, 14–19, 100, 127, 149, 232–233, 260, 261, 263, 326, 335–336, 372, 375, 388, 390, 391, 411–421, 424, 428, 436–438, 441, 442, 444, 446–450

Arc flow, 232, 390

Arrival penalty, 17, 423, 437, 438, 448

Arrow, K.J., 6, 9, 120, 122–124, 317, 344

Arrow sufficiency theorem, 120, 122, 145, 339, 340

Aspatial monopolistic firm, 353, 362–367

B

Bellman, R.E., 114–116

Beltrami identity, 96–97, 102

Bernoulli, J., 3, 4

Brachistochrone problem, 2, 4, 96, 100–103, 144

C

Calculus of variations, 1–4, 79–145, 147, 150, 163, 169, 185, 186

Cell transmission, 415

Coerciveness, 202, 216

Contraction, 199, 203, 415, 420

Contraction mapping theorem, 156, 285

Control, 14, 20, 23–27, 69–78, 103–106, 108–110, 112, 114, 115, 117–122, 124, 126–131, 135, 138, 142, 144, 145, 163, 165, 169–171, 174–178, 183, 190, 209, 213, 220, 257, 268, 271–278, 281–284, 287–289, 292, 294, 296, 297, 299, 301, 303–311, 314–319, 325–329, 332–337, 340, 347, 349, 354–357, 359–361, 363–365, 367–369, 371, 373, 376, 377, 379–381, 383–387, 389, 397–400, 402–404, 407, 414, 415, 418, 420, 421, 430–434, 438–441, 445, 446, 448, 452, 465, 469, 471–473, 480, 485–487

Control variable, 3, 6, 14, 19, 70, 71, 73, 103–105, 112, 117, 120–122, 127, 131, 165, 174, 315, 316, 325, 327, 334–336, 340–342, 349, 368, 373, 385, 389, 407, 414, 418, 420, 421

Convergence, 131, 149, 152–154, 192–199, 202–204, 241, 242, 254, 257–259, 283, 284, 287–289, 298, 300, 305, 408, 432, 439, 445, 448, 454, 464, 476, 488

Convex functions, 34, 53–58, 224, 236, 460

Convexity, 33, 34, 52–62, 122, 126, 157–158, 197, 199, 202, 230, 231, 236, 251, 279, 283, 291–294, 321, 326, 461

Criterion, 3, 8, 13–14, 33, 81, 103, 104, 113, 116, 117, 128, 174, 180, 182, 190, 207, 274, 294, 301, 340, 343, 348–349, 355, 370, 461, 462

D
Day-to-day dynamic, 411
Demand management, 457
Differential game, 1, 21, 33, 272, 282, 457
Differential variational inequality, 2, 22, 26–27, 271–277, 376, 382, 383, 401–404, 408, 411, 412, 429, 431, 432, 437–440, 445, 453, 464–466
Discretization, 130, 209, 289, 349, 370, 371, 384, 385, 387, 464, 476
Distributions, 23, 353, 354, 367, 369, 376, 377, 405, 453
Dual variable, 72, 117, 134, 246, 274, 301, 302, 331, 358–360, 363, 366, 376, 381, 430, 431, 433, 435, 440, 473
Dynamic game, 1, 20–21, 30, 219, 267, 273, 279, 353
Dynamic oligopolistic spatial competition, 353, 376–395
Dynamic programming, 21, 114–117
Dynamic traffic assignment (DTA), 1, 174, 273, 411, 413, 415, 453
Dynamic user equilibrium, 2, 21–22, 30, 411–454

E
Economic growth, 2, 4–6, 10, 11, 313–351
Effective path delays, 412, 423, 428, 429, 432, 445, 451, 453
Entrance flow, 413–415, 417, 418, 421
Euler, 3, 4, 91
Euler-Lagrange equation, 79, 85–87, 93–99, 101, 147, 148, 185–186
Existence, uniqueness, 21, 235–237, 250, 408
Exit flow, 19, 413–418, 421, 423, 438, 448–450, 452
Exit time function, 15–17, 412, 416–419, 437
Extremal, 3, 23–26, 29–30, 81, 87, 96, 118, 220, 237, 297, 326, 332, 335, 355, 367–369, 380, 384, 396–401, 428, 471, 485–487

F
Factor, 397, 398, 400, 401, 406, 458
Feasible set, 237, 382
Fixed endpoint(s), 95
Fixed inventories, 458–459
Fixed point problem, 20, 177, 221, 222, 228, 231, 235, 255–257, 268, 282, 293, 302, 303, 432, 434, 445, 452, 488
Flow conservation, 21, 23, 232, 369, 377, 379, 420, 428–429, 435, 440
Flow propagation constraint, 412, 414, 415, 418–421, 424, 436, 438, 446, 458
Fréchet derivative, 162–163
Free endpoint(s), 95, 99
Fritz John conditions, 38–39, 41–50
Functional, 3, 4, 13–14, 24, 25, 30, 81–82, 87, 92, 96, 100, 103, 105, 109, 114–116, 120, 128, 147–155, 157–165, 185–187, 192, 194, 198, 199, 202, 207, 216, 277, 278, 281, 282, 294, 295, 348–349, 354, 368, 369, 373, 378, 379, 383, 461

G
Game, 1, 2, 20–26, 219–221, 267, 272, 273, 282, 380, 382, 383, 401, 408, 411, 458, 463, 467–481, 491
Gap function, 220, 248–255, 263, 289–299, 464–466
Gap function algorithm, 464
Gâteaux derivative, 147, 159–162
G-derivative, 160, 162–169, 172, 174, 178–180, 185, 187, 189, 198, 216, 294, 295, 461
Generalized Nash equilibrium
  dynamic, 22, 26, 219, 267, 277, 281, 282, 413, 458
  static, 2, 20, 21, 219, 267, 272
Global minimum (maximum), 34, 35, 55, 58–61, 68, 121, 164, 188, 224, 340
Global optimality, 33, 52, 58, 68, 223,
Gradient, 37, 40, 50, 51, 85, 149, 159–162, 165, 182, 185, 190–192, 195, 198–201, 207, 209–214, 217, 224, 225, 231, 236, 238, 250, 252–254, 267, 286, 287, 294, 295, 297, 298, 305, 326, 328–335, 373, 375, 376, 402–404, 452, 462, 466, 476, 481
Gradient projection algorithm, 200–202, 209–214, 287, 373

H
Hahn-Banach theorem, 159–162
Hamiltonian, 74, 75, 77, 106, 112, 116, 117, 119–122, 125, 127, 129, 139, 169, 173, 174, 183, 191, 215, 274, 279,

294, 301, 303, 326, 328, 330, 333–335, 340–343, 356, 359, 363, 364, 376, 381, 383, 401–404, 408, 415, 430, 432, 440, 473, 486
Hamilton-Jacobi-Bellman partial differential equation, 21
Hilbert space, 147, 148, 155, 156, 160, 161, 163, 165, 185–190, 198–200, 267, 272, 286, 289–298, 411, 421

I
Incidence matrix, 28, 399, 429, 470, 481, 483
Inner product, 150, 160, 161, 271, 272, 429
Integer program, 37
Interior point, 149, 159
Interior point constraint, 113–114
Inventory, 2, 24, 25, 31, 273, 353–355, 357–359, 366, 368, 369, 371, 372, 374–376, 378–381, 387, 389, 390, 392, 397–401, 406, 407, 458, 459

K
Kuhn-Tucker conditions
  finite, 77, 186, 187, 224, 276
  infinite, 147, 148, 186, 187

L
Lagrange, 3, 354
Lagrangean, 39, 72, 106, 116, 118
Lagrange multiplier, 41–43, 96, 297
Lemke's method, 259
Linear complementarity, 256
Linear program (LP), 35, 36, 62–64
Linear quadratic problem (LQP), 138–144
Line integral, 230, 239, 240, 242, 243, 428,
Local minimum (maximum), 35, 38, 42, 43, 48, 49, 55, 56, 118, 188, 451

M
Mangasarian, O.L., 48, 120, 121
Mangasarian sufficiency theorem,121, 174
Mathematical program, 1, 3, 30, 33–37, 39, 41, 42, 44, 46–49, 52, 55, 59, 62, 66, 68, 69, 71, 72, 77, 79, 103, 118, 120, 130, 147–217, 220–226, 228, 242, 245, 250–252, 257, 268, 271–273, 289, 313, 370, 373, 383, 440, 459
Maximum principle, 3, 70, 325, 326, 329, 330, 334, 339, 340, 351, 382, 401–403, 409, 473, 486
Measurable, 154, 422, 424,
Minimum principle, 3, 34, 37, 74–77, 79, 108, 112, 117, 119, 122–125, 127, 130, 135, 139, 169, 274, 275, 279, 280, 403–404, 431, 433, 440, 441
Monopoly, 353, 362–364
Monotone function, 236, 439
Monotonicity, 236, 445
Multiple time scales, 23, 377

N
Nash
  equilibrium
  dynamic, 22, 26–27, 219, 277, 281–282, 376, 458–459, 474–475
  static, 2, 20, 21, 219, 267, 272
  game
  dynamic, 1, 20, 22, 219, 267, 273, 277, 282, 486
  static, 20, 219, 267, 272
Necessary conditions, 20, 22, 33–34, 37–46, 71–74, 82, 86–93, 96, 105–114, 129, 132, 136, 138, 144, 145, 147, 148, 160, 163–165, 169–174, 177–182, 186, 224–226, 267, 268, 274–275, 300–303, 335, 337, 339, 343, 363–364, 370, 382, 383, 436, 439–442, 473, 487
Nested operator, 437–438
Network, 2, 14, 15, 18, 19, 21–28, 220, 231–232, 260–263, 353, 367–377, 380, 383–384, 386, 395, 405, 409, 411–418, 420, 423, 424, 428, 436, 440, 445, 447–448, 454, 457–458, 481–491
Network design, 457
Node, 14–15, 23–25, 232, 260, 353, 367, 368, 371–372, 376–378, 381, 386–387, 389–390, 395–400, 405–406, 419, 447, 454, 488
Nonextremal problem, 219, 221–237
Nonlinear complementarity problem (NCP), 20, 219, 221, 223, 227, 228, 230, 233, 259, 260, 268, 276–277, 289, 389
Nonlinear program, 1, 21, 33–79, 103, 147, 148, 217, 225–226, 230, 241, 253, 258, 263, 351, 382, 465, 473
Norm, 125, 148, 150–152, 155–157, 198–199, 210, 222, 223, 228, 253, 257, 258, 282, 303, 432

O
Oligopolistic competition, 2, 23, 353, 377, 395, 468, 475
Oligopoly, 27, 353–409, 458, 481
Open loop, 267, 277, 353, 354
Optimal control, 1–3, 11, 14, 20, 30, 33, 76, 79–145, 147, 148, 150, 163, 165, 183, 190, 209, 284, 313, 342, 343, 347, 434, 440, 441
Optimal control problem, 3, 6, 7, 20, 22, 25, 29, 33, 69–71, 74, 75, 77–79, 104–105, 109, 114–117, 120, 121, 124, 127, 130–131, 138, 144, 147, 165, 169–171, 174, 176, 177, 183, 209, 216, 271, 274, 278, 281, 283, 289, 301, 303–304, 313–317, 325, 355, 356, 359, 364, 369, 379, 381, 383, 400, 430, 432, 434, 439–441, 445, 471–473, 485–487
Optimal growth, 6–7, 313, 315–317
Optimality, 31, 33–35, 37, 39, 43, 46–48, 52, 58, 68, 71, 77, 79, 80, 93, 97, 108, 114, 120, 122, 123, 134, 138–140, 147–148, 158–160, 174, 182, 187, 188, 215, 223, 224, 230, 291, 304, 313, 351, 435
Origin-destination (OD) pair, 15, 18, 23–24, 232, 367, 377, 398, 414, 420, 447, 454, 489
Overbooking, 457, 458, 485

P
Path delay operator, 17–18, 411–413, 415, 423, 428, 429, 432, 437, 445
Path flow, 18, 233, 260, 353, 386, 390, 391, 394, 412, 418, 421, 423, 448–451
Path variable, 386, 388
Penalty function, 190, 204–206, 214–216, 423
Perfect competition, 353–358, 377
Pontryagin, 3, 70, 74, 75, 79, 120, 147, 275, 325, 326, 334, 339
Population, 5–11, 13, 314, 315, 317, 321, 343–348
Price, 27–30, 72, 116, 134, 353–355, 362, 364, 381, 390, 391, 395, 396, 441, 457–459, 463–470, 476–479, 481–486, 491, 492,
Price taking, 354–362
Pricing, 2, 21, 27–30, 106, 111, 273, 457–492
Principle of optimality (POO), 114
Producer, 376, 395–402, 405–407
Production function unit cost, 23–24, 232, 368, 372, 378–380, 388, 390, 397, 400–401, 406, 471
Projection operator, 12, 125, 257, 284, 293, 348, 434–435
Public investment, 2, 7–14, 318–320, 323, 325, 327–334, 336, 343–350

R
Ramsey, F.P., 2, 5–7, 314–317
Refund(s), 27–30, 481, 482, 485, 486, 490
Retailer, 27, 395–399, 402–408, 457
Revenue management (RM), 1, 2, 21, 27–31, 273, 457–492
Ricatti equation, 141
Riesz representation theorem, 156, 157, 160, 185

S
Separable function, 240, 241
Service cost, 28–30, 482, 485, 486, 490
Shipping, 2, 23–25, 367–370, 377, 379, 398, 408
Singular control, 20, 76, 77, 119–120, 127–130, 313, 326, 327, 333, 335, 337, 341, 351, 357, 360, 361, 479
Sobolev space, 23, 161, 170, 175, 271, 300, 354, 367, 377
Space of square integrable functions, 23, 161, 170, 271, 300, 354, 367, 377, 421
Spatial competition, 353, 376–395
State, 3, 6, 11, 14, 17, 19–21, 23, 25, 34–37, 41, 42, 44, 47–50, 53, 79, 83, 84, 96, 100, 104–108, 110, 111, 114, 148, 151, 156, 159, 163, 165, 166, 168–171, 174–177, 219, 224, 231–233, 237, 241, 242, 248, 250, 255, 268, 271–278, 281, 282, 286–289, 291, 294, 297, 315–317, 328, 332, 339, 340, 342, 347, 354, 360, 362, 367, 369–371, 373, 412, 418–423, 428–430, 438–441, 446, 448, 460, 462, 463, 469, 472–474, 476
State-dependent time shift, 174, 303, 305, 307, 440, 441, 445
State operator, 166, 169, 170, 175, 268, 271, 272, 377, 380, 408, 438
State variable, 3, 6, 12, 14, 21, 23, 25, 28, 70, 71, 74, 76, 104, 105, 108, 116, 117, 120, 122, 128, 135, 165, 271, 272, 274, 286, 288, 301, 306, 315–317,

340, 342, 348, 349, 360, 367, 369, 371, 377, 380, 385, 387, 420, 421, 423, 430, 438, 440, 441, 446, 464, 473, 482
Steepest descent algorithm, 190–198, 206–210, 217
Successive linearization, 256, 259, 385
Sufficient conditions, 1, 52–60, 80, 93, 124, 145, 199, 223, 225, 226, 228, 230, 231, 249, 251, 279, 408, 462, 463, 499
Supply chain, 1, 31, 174, 273, 353–409
Symmetry restriction, 242
System optimal flow, 18, 415

T
Technological change, 8, 12, 343, 347–348
Terminal constraints, 109–111, 113, 144, 170, 271, 273, 274, 301
Terminal cost, 109, 111, 144, 169, 216, 380
Terminal time, 24, 29, 81, 105, 106, 109–112, 117, 120–122, 134, 138, 144, 165, 169, 293, 297, 321, 328, 330, 353–355, 358, 359, 365, 368, 372, 376, 378–381, 387, 390, 398, 399, 408, 420, 440, 448, 458, 465, 469, 473, 474
Time scale, 23, 377, 411, 451
Time shift, 20, 148, 174–186, 216, 268, 298–309, 311, 370, 438, 440, 441, 445–446
Topological vector space, 33, 147, 149–157, 163, 268–270
Total cost, 395
Transversality condition, 74, 76, 108, 119, 125, 129, 134, 169, 172, 178, 274, 275, 301, 326, 335, 338, 361, 430, 433, 445, 487
Two-point boundary value problem, 74, 101, 106, 119, 130–132, 134, 136, 139, 170, 175, 209, 271, 360, 365, 429, 463, 464, 474

U
User equilibrium, 2, 21–22, 30, 228, 231–235, 260–263, 411–454
User optimal flow, 232

V
Variation, 1–4, 20–22, 26, 27, 30, 79–144, 219–263, 351, 376, 382–385, 389, 401, 404, 408, 409, 411–413, 423–425, 427–432, 434, 436–440, 445, 447, 453
Variational inequality, 2, 21, 22, 26, 27, 30, 109, 147, 164, 165, 169, 171, 174, 177, 178, 182–184, 199, 200, 202, 220, 221, 223, 224, 226–231, 234–238, 241–243, 245, 246, 248, 255–259, 261–263, 267, 271–274, 279, 280, 282, 283, 289, 293, 294, 300–305, 310, 376, 382–385, 389, 401–405, 408, 411–413, 423–425, 427, 429–432, 436, 437, 439, 440, 445, 453, 460–466, 473–475, 486–490
  finite, 27, 219–263, 290, 292–293, 304, 382, 389, 404, 408, 428, 481
  infinite, 21, 30, 147, 174, 267, 412, 423, 425, 460, 463, 464
Vector space, 33, 37, 69, 105, 138, 147, 149–152, 154–156, 158, 160, 163, 185, 186, 202, 268–270

W
Weierstrass theorem, 89, 91, 94, 163
Within-day dynamics, 411